The Argo simulation - I. Quenching of massive galaxies at high redshift as a result of cosmological starvation

Feldmann, R ; Mayer, L

Abstract: Observations show a prevalence of high-redshift galaxies with large stellar masses and predominantly passive stellar populations. A variety of processes have been suggested that could reduce the star formation in such galaxies to observed levels, including quasar mode feedback, virial shock heating, or galactic winds driven by stellar feedback. However, the main quenching mechanisms have yet to be identified. Here we study the origin of star formation quenching using Argo, a cosmological, hydrodynamical zoom-in simulation that follows the evolution of a massive galaxy at z 2. This simulation adopts the same subgrid recipes of the Eris simulations, which have been shown to form realistic disc galaxies, and, in one version, adopts also a mass and spatial resolution identical to Eris. The resulting galaxy has properties consistent with those of observed, massive (M* ˜ 1011 M) galaxies at z ˜ 2 and with abundance matching predictions. Our models do not include active galactic nuclei (AGN) feedback indicating that supermassive black holes likely play a subordinate role in determining masses and sizes of massive galaxies at high-z. The specific star formation rate (sSFR) of the simulated galaxy matches the observed M*-sSFR relation at early times. This period of smooth stellar mass growth comes to a sudden halt at z = 3.5 when the sSFR drops by almost an order of magnitude within a few hundred Myr. The suppression is initiated by a levelling off and a subsequent reduction of the cool gas accretion rate on to the galaxy, and not by feedback processes. This 'cosmological starvation' occurs as the parent dark matter halo switches from a fast collapsing mode to a slow accretion mode. Additional mechanisms, such as perhaps radio mode feedback from an AGN, are needed to quench any residual star formation of the galaxy and to maintain a low sSFR until the present time.

DOI: https://doi.org/10.1093/mnras/stu2207

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **446,** 1939–1956 (2015) doi:10.1093/mnras/stu2207

The Argo simulation – I. Quenching of massive galaxies at high redshift as a result of cosmological starvation

Robert Feldmann[1]★† and Lucio Mayer[2]

[1]*Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA*
[2]*Center for Theoretical Astrophysics and Cosmology, Institute for Computational Science, University of Zürich, CH-8057 Zürich, Switzerland*

Accepted 2014 October 18. Received 2014 October 3; in original form 2014 April 18

ABSTRACT

Observations show a prevalence of high-redshift galaxies with large stellar masses and predominantly passive stellar populations. A variety of processes have been suggested that could reduce the star formation in such galaxies to observed levels, including quasar mode feedback, virial shock heating, or galactic winds driven by stellar feedback. However, the main quenching mechanisms have yet to be identified. Here we study the origin of star formation quenching using Argo, a cosmological, hydrodynamical zoom-in simulation that follows the evolution of a massive galaxy at $z \geq 2$. This simulation adopts the same subgrid recipes of the Eris simulations, which have been shown to form realistic disc galaxies, and, in one version, adopts also a mass and spatial resolution identical to Eris. The resulting galaxy has properties consistent with those of observed, massive ($M_* \sim 10^{11}\,\mathrm{M}_\odot$) galaxies at $z \sim 2$ and with abundance matching predictions. Our models do not include active galactic nuclei (AGN) feedback indicating that supermassive black holes likely play a subordinate role in determining masses and sizes of massive galaxies at high-z. The specific star formation rate (sSFR) of the simulated galaxy matches the observed M_*–sSFR relation at early times. This period of smooth stellar mass growth comes to a sudden halt at $z = 3.5$ when the sSFR drops by almost an order of magnitude within a few hundred Myr. The suppression is initiated by a levelling off and a subsequent reduction of the cool gas accretion rate on to the galaxy, and not by feedback processes. This 'cosmological starvation' occurs as the parent dark matter halo switches from a fast collapsing mode to a slow accretion mode. Additional mechanisms, such as perhaps radio mode feedback from an AGN, are needed to quench any residual star formation of the galaxy and to maintain a low sSFR until the present time.

Key words: galaxies: evolution – galaxies: high-redshift – galaxies: star formation.

1 INTRODUCTION

A large number of massive galaxies at $z \sim 2$ are dominated by a passively evolving stellar population with little ongoing star formation (e.g. Franx et al. 2003; Cimatti et al. 2004; Förster Schreiber et al. 2004; Daddi et al. 2005; van Dokkum et al. 2006; Arnouts et al. 2007; Wuyts et al. 2007; Cimatti et al. 2008; Kriek et al. 2008; Brammer et al. 2009; Williams et al. 2009; Ilbert et al. 2010; McCracken et al. 2010; Onodera et al. 2012; Bezanson et al. 2013; Whitaker et al. 2013; Lundgren et al. 2014; Magnelli et al. 2014).

The fraction of such galaxies is 20–30 per cent for galaxies with stellar masses $\sim$2–5 $\times$ 10^{10} $\mathrm{M}_\odot$ and raises to $\sim$50 per cent for galaxies with masses above $\sim 10^{11}$ $\mathrm{M}_\odot$ (e.g. Brammer et al. 2011; Whitaker et al. 2011; Muzzin et al. 2013; Tal et al. 2014). These galaxies are called quiescent, a reference to their weak or absent star formation activity, although the actual operational definition is typically based on $U - V$ and $V - J$ rest-frame colours, see Section 3. So far, no consensus has been reached on the nature of the physical processes responsible for suppressing ('quenching') star formation in massive galaxies at high redshift.

Empirical studies show that out to at least $z \sim 1$ the likelihood of a galaxy being quenched is correlated both with its stellar mass ('mass quenching') and with the environmental density of neighbouring galaxies ('environmental quenching'), see Peng et al. (2010). The latter quenching channel is primarily affecting satellite galaxies (e.g. Peng et al. 2012; Wetzel et al. 2013), while the quiescent fraction of central galaxies appears to be not strongly correlated with current environmental overdensity (Kovač et al. 2014, and references therein). Environmental quenching is likely related to processes preferentially occurring in galaxy groups and clusters,

★Hubble fellow.
†E-mail: feldmann@berkeley.edu

such as ram-pressure stripping of the interstellar medium (ISM; Gunn & Gott 1972; Abadi, Moore & Bower 1999), the reduction of gas accretion on to satellites ('starvation'; Larson, Tinsley & Caldwell 1980; Balogh, Navarro & Morris 2000; Kawata & Mulchaey 2008; McCarthy et al. 2008; van den Bosch et al. 2008; Feldmann, Carollo & Mayer 2011; Bahé et al. 2013), or frequent galaxy interactions ('harassment'; Farouki & Shapiro 1981; Moore et al. 1996).

Most mass quenching mechanisms discussed in the literature either prevent or decrease the accretion of gas on to galaxies or eject gas from galaxies. For instance, the growth of a stable virial shock in massive haloes reduces the accretion of relatively cool, not strongly shock-heated, gas on to galaxies (Birnboim & Dekel 2003; Kereš et al. 2005; Cattaneo et al. 2006; Dekel & Birnboim 2006). In addition, cooling of shock-heated gas might be counteracted by gravitational heating from infalling satellite galaxies or gas clumps (Dekel & Birnboim 2008; Khochfar & Ostriker 2008; Johansson, Naab & Ostriker 2009, 2012).

Other quenching channels are related to active galactic nuclei (AGN) and to star formation in galaxies. Major mergers of gas-rich disc galaxies at high redshift may ignite powerful starbursts (Dekel & Silk 1986; Murray, Quataert & Thompson 2005) and result in quasar activity (Sanders et al. 1988; Scannapieco & Oh 2004; Di Matteo, Springel & Hernquist 2005; Springel, Di Matteo & Hernquist 2005a,b; Hopkins et al. 2006; McCarthy et al. 2011) that may launch outflows and remove large amounts of gas from galaxies. AGN can also operate in a jet-powered *radio mode* in which they counteract the cooling of gas from the hot halo surrounding galaxies with low intensity mechanical heating (e.g. Bower et al. 2006; Croton et al. 2006; Sijacki et al. 2007).

Ejective feedback originating in quasar activity could explain why the entropy in galaxy clusters is close to the value needed to offset gas cooling (Scannapieco & Oh 2004). It also offers a physical basis for the co-evolution of black hole mass and the bulge mass of the galaxy host (Magorrian et al. 1998; Marconi & Hunt 2003; Häring & Rix 2004) and for the correlation between quasar activity and star formation (Boyle & Terlevich 1998; Kauffmann et al. 2003b). However, it is still an open question whether quasar-driven outflows triggered by galaxy interactions or mergers actually quench star formation in massive, high-redshift galaxies (see Alexander & Hickox 2012, for a recent review). In fact, at high redshift there is little evidence for a link between merging and quasar activity (Schawinski et al. 2011; Kocevski et al. 2012), except at the highest quasar luminosities (Treister et al. 2012), or between quasar activity and star formation (Goulding et al. 2014).

Galaxy evolution simulations that include quasar mode feedback have not yet converged on a conclusive answer either, primarily because there exists a variety of choices as to which physical processes are modelled and how they are implemented (e.g. Debuhr et al. 2010; Ostriker et al. 2010; Novak, Ostriker & Ciotti 2011; Power, Nayakshin & King 2011; Faucher-Giguère & Quataert 2012; Newton & Kay 2013; Wurster & Thacker 2013; Gabor & Bournaud 2014).

Numerical simulations agree, however, that radio mode feedback can reduce the $z = 0$ stellar masses of central galaxies in clusters, that it can result in more realistic galaxy colours and also better agreement with observed X-ray scaling relations (Sijacki & Springel 2006; Khalatyan et al. 2008; Puchwein, Sijacki & Springel 2008; Dubois et al. 2010, 2011, 2013; McCarthy et al. 2010; Teyssier et al. 2011; Ragone-Figueroa et al. 2013; Planelles et al. 2014). However, this feedback channel primarily affects the hot gas halo surrounding galaxies, and typically does not launch large-scale winds involving the ISM of the galaxy. Hence, radio mode feedback appears to be an unlikely candidate for triggering a rapid star formation quenching in high-redshift galaxies, but it may play an important role in maintaining low star formation activity in already quenched galaxies (e.g. Best et al. 2005; Croton et al. 2006).

Hence, we are left with two related, but distinct, questions. First, why do high-redshift galaxies leave the star-forming sequence (Daddi et al. 2007; Elbaz et al. 2007, 2011; Noeske et al. 2007)? We will show that massive, high-redshift galaxies can only remain on the star-forming sequence for so long before the finite supply of gas limits their star formation activity. Specifically, during an early collapse phase, gas accretion rates and specific star formation rates (sSFRs) are high and massive galaxies form and assemble much of their stellar mass. This is followed by a cosmological starvation phase in which accretion rates, and subsequently star formation activity, level off and eventually decline. Secondly, why do galaxies shut down their star formation almost completely instead of maintaining a low, but non-negligible star formation activity? We will argue that a combination of stellar feedback and, potentially, AGN radio mode feedback, coupled with gravitational heating is required to complete the transition from the star-forming sequence to the quiescent galaxy population.

Our results are based on a cosmological, zoom-in simulation of a massive, high-redshift galaxy. The simulated galaxy resides at the centre of a halo ($M_{\rm vir} \sim 10^{13}$ M$_\odot$ at $z = 0$) that should harbour common (i.e. of intermediate stellar mass $\sim$1–3 $\times$ 10^{11} M$_\odot$) quiescent galaxies in the local Universe. The halo is located in a typical, mildly overdense region that does not contain a more massive halo. Hence, our findings likely apply to a majority of massive, quiescent, central galaxies.

Despite their importance and ubiquity, galaxies in $\sim 10^{13}$ M$_\odot$ haloes have only recently been targeted by cosmological, zoom-in simulations (Kawata & Mulchaey 2008; Feldmann et al. 2010, 2011; Oser et al. 2010, 2012; Cen 2014). One of the reasons is the numerical challenge of resolving galaxies in such massive haloes. In addition, physical models are often tuned based on simulations of lower mass, star-forming galaxies (e.g. Stinson et al. 2006). Hence, it is not obvious that the same models are adequate to simulate massive, quiescent galaxies. Fortunately, we can gauge the realism of our numerical approach even before starting the Argo simulation. In particular, simulations run with the same code, with similar methodology, and at a comparable resolution produce dwarf galaxies (Shen et al. 2014), Milky Way like galaxies (e.g. the Eris simulation; Guedes et al. 2011), and massive galaxies (Feldmann et al. 2010) with reasonably realistic properties. The same numerical approach is also able to reproduce the enrichment level of the circumgalactic gas around high-redshift galaxies (Shen et al. 2013).

The paper is organized as follows. We outline the set-up of the Argo simulation and the strategy of the data analysis in Section 2. We present our first results in Section 3, where we compare the properties of the simulated galaxy with those of observed high-redshift galaxies. In Section 4 we then analyse the star formation history of the simulated galaxy finding an onset of quenching at $z \sim 3.5$. We identify cosmological starvation as the cause of the star formation suppression in Section 5. We highlight potential caveats that may affect the conclusion drawn from our work in Section 6. We compare our results to previous theoretical works in Section 7. Finally, in Section 8 we end this paper with a summary of our main findings and our conclusions.

Table 1. Overview of the individual runs performed as part of the Argo project. The first four columns provide the label and the masses of gas, star, and DM particles in the zoom-in region of the run. The next two columns are the gravitational spline softening lengths of baryonic (i.e. gas and star) particles and of DM particles in proper units. The penultimate column shows the star formation threshold. The final column indicates the redshift at which the run is stopped. Runs MR1 and MR2 start from the same initial conditions and differ only in the adopted star formation threshold.

Run	$m_{\rm gas}$ ($\mathrm{M}_\odot\,h^{-1}$)	$m_{\rm star}$ ($\mathrm{M}_\odot\,h^{-1}$)	$m_{\rm DM}$ ($\mathrm{M}_\odot\,h^{-1}$)	$\epsilon_{\rm bar}$ (pc h^{-1})	$\epsilon_{\rm DM}$ (pc h^{-1})	$n_{\rm SF}$ ($\mathrm{m_H\,cm^{-3}}$)	$z_{\rm end}$
LR[a]	9.9×10^5	2.9×10^5	4.7×10^6	219	365	0.1	1
MR1	1.2×10^5	3.7×10^4	5.8×10^5	109	183	0.1	2
MR2	1.2×10^5	3.7×10^4	5.8×10^5	109	183	5	2
HR	1.5×10^4	4.6×10^3	5.8×10^5	88	183	5	3.4

[a]We carried out at this resolution: the default LR run with our fiducial physics model, the LR-noFB run without energetic feedback from SNe, the LR-noFB($z < 4$) run without SN feedback after $z = 4$, and the LR-noML($z < 3$) run without stellar mass-loss after $z = 3$.

2 DETAILS OF THE SIMULATION AND THE DATA ANALYSIS

In this paper we present Argo, a cosmological, hydrodynamical simulation of a group-sized dark matter (DM) halo and its gaseous and stellar content. The virial mass of the selected DM halo is $\sim 3 \times 10^{12}\,\mathrm{M}_\odot$ at $z = 2$ and $\sim 2 \times 10^{13}\,\mathrm{M}_\odot$ at $z = 0$. Here and throughout the paper, the virial mass of a halo refers to the mass enclosed within a spherical volume centred on the density peak of the halo with a mean matter density of 180 times the background density of the Universe.

The Argo simulation consists of a set of individual runs, with different choices for the numerical resolution or model parameters, of the same physical system. The four main runs (low resolution, LR, medium resolution 1, MR1, medium resolution 2, MR2, and high resolution, HR) of the Argo project are summarized in Table 1. The simulation is a follow-up of the *G*2 simulation introduced in Feldmann et al. (2010). However, some of the runs have significantly improved numerical resolution and updated subgrid model parameters that match those of the 'Eris' simulation (Guedes et al. 2011).

We select the DM halo with a target $z = 0$ virial mass from a 512^3 *N*-body simulation of a $(123\ \mathrm{Mpc})^3$ cosmological volume (Hahn et al. 2007). We ensure that there is not a more massive halo within 7 Mpc and that the selected halo is not sitting in a void. The average overdensity is 1.4 when measured within in a sphere of 7 Mpc radius centred on the selected halo. No additional selection criteria are used.

To increase the numerical resolution at and around the halo we identify the Lagrangian patch in the initial conditions ($z_{\rm init} = 41.5$) that contains all particles that enter a sphere of radius $R(z) = 2 \times R_{\rm vir}(z = 0)/(1 + z)$ around the most massive progenitor of the halo at any $z \geq z_{\rm m}$. We use $z_{\rm m} = 2$ for the runs MR1, MR2, and HR and $z_{\rm m} = 0$ for run LR. We use GRAFIC-2 (Bertschinger 2001) to add higher frequency density fluctuations to the original DM simulation and to obtain DM and smoothed particle hydrodynamics (SPH) particle positions and velocities in the Zeldovich approximation. We refine the Lagrangian patch to the chosen DM resolution, see Table 1, and embed it into spherical shells of decreasing resolution. For each DM particle in the HR Lagrangian patch of the LR and MR runs we add one SPH particle while ensuring the proper baryonic power spectrum of the gas particles. We create initial conditions for the HR run based on the initial condition of the MR runs using a basic SPH particle splitting scheme. In particular, we split each SPH particle into eight lower mass SPH particles and offset their positions relative to the position of the original SPH particle.

The initial power spectrum of density fluctuations is generated with LINGER (Bertschinger 1995) and is compatible with *Wilkinson Microwave Anisotropy Probe* 3 cosmological parameters (Spergel et al. 2007). Specifically, we use $\Omega_{\rm m} = 0.24$, $\Omega_\Lambda = 0.76$, $\Omega_{\rm b} = 0.04185$, $H_0 = 73\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\sigma_8 = 0.77$, and $n = 0.96$.

The Argo simulation is run with the parallel TreeSPH code GASOLINE (Wadsley, Stadel & Quinn 2004). GASOLINE is based on the parallel, multiple time-stepping *N*-body code PKDGRAV (Stadel 2001). We keep the gravitational softening length fixed in proper coordinates for $z < 9$ and fixed in comoving coordinates for $z \geq 9$. We provide the mass and force resolution of the various runs in Table 1.

In addition to gravitational and hydrodynamical processes, the Argo simulation includes heating caused by a spatially uniform, redshift dependent UV radiation background (Haardt & Madau 1996; Guedes et al. 2011), (optically thin) radiative cooling for a primordial gas composition (Wadsley et al. 2004), and subgrid models for star formation and stellar feedback (Stinson et al. 2006). Star particles (representing single stellar populations) form in a probabilistic fashion out of gas that is part of a convergent flow, has temperatures below 3×10^4 K, and spatial densities above $n_{\rm SF}$, see Table 1. The star formation rate (SFR) proceeds at a rate $\dot{\rho}_* = \epsilon_{\rm SF}\rho_{\rm g}/t_{\rm dyn} \propto \rho_{\rm g}^{1.5}$. The star formation efficiency per free-fall time, $\epsilon_{\rm SF}$, is 5 per cent. Subgrid stellar winds return a moderate fraction (~40 per cent within a few Gyr) of the mass bound in star particles back to the gas.

The number of Type II and Ia supernovae (SNe) that occur during a given time step is computed based on a Kroupa, Tout & Gilmore (1993) initial stellar mass function (IMF) and the stellar tracks by Raiteri, Villata & Navarro (1996). Each SN injects metals and a thermal energy of 8×10^{50} erg into the neighbouring SPH particles. We use the analytic blast wave scenario of McKee & Ostriker (1977) to estimate the radius and cooling time of the unresolved Type II SN blast waves. We suspend cooling within the SPH kernel surrounding the star particle for a time corresponding to the end of the snowplough phase of the Type II SN blast wave. For typical ISM conditions the shutoff time is of the order of a few $\sim 10^5$–10^6 yr. Cooling is not suspended for Type Ia SN.

The LR run is continued to $z = 1$, the MR1 and MR2 runs to $z = 2$, and the HR run down to $z = 3.4$. We store snapshots every ~70–150 Myr.

We identify DM haloes and subhaloes with the AMIGA Halo Finder (AHF; Gill, Knebe & Gibson 2004; Knollmann & Knebe 2009). We run AHF on the total matter field generated by

star, gas, and DM particles. In this paper we focus on the most massive galaxy in the HR region and on its DM halo. We use the merger tree tool provided by AHF to trace this galaxy from one snapshot to the next. In addition, we use the merger tree tool to identify subhaloes/satellite galaxies surrounding the main galaxy. Sometimes we find it useful to remove these subhaloes/satellite galaxies in our analysis, see below. We found that AHF sometimes identifies substructures within the main galaxy that are, e.g. star-forming regions or density enhancements in spiral arms. Hence, we exclude substructures within a 2 kpc physical radius from the centre of the galaxy from our substructure list.

Reported stellar masses, optical and infrared magnitudes, and SFRs of the simulated galaxy do not include contributions from satellite galaxies and are measured within a spherical radius of 12 kpc, unless stated otherwise. We checked that increasing the radius to 20 kpc changes these properties only at the few per cent level. This is probably not surprising given that the half-mass radius of the central galaxy is $\lesssim$1 kpc at $z \geq 2$, see Section 3. For the $z = 1$ snapshot (LR run only), we use a somewhat larger radius of 20 kpc to account for the increased size of the galaxy at late times. Half-mass radii are 2D apertures including exactly half the stellar mass (as defined above) of a galaxy. As projected radii depend on viewing direction (typical variations are of the order of 10 per cent), we quote mean half-mass radii averaged over 25 different lines of sight, including 20 random lines of sight, the x, y, and z projection, and a face-on and an edge-on projection. SFRs at any given time refer to the stellar mass formed within the past 100 Myr, unless otherwise noted.

We use the stellar population synthesis model of Bruzual & Charlot (2003) to estimate the luminosities associated with each individual stellar particle. In particular, the model interprets each stellar particle as a single stellar population with a Chabrier (2003) IMF and with the mass, age, and metallicity as recorded by the stellar particle. Whenever necessary we interpolate logarithmically between the provided grid of metallicities and stellar population ages. We compute both rest-frame magnitudes (in the AB system) in standard Bessel U and V filter bands and in the NSFCam J filter band,[1] as well as apparent K-band[2] magnitudes. We add a moderate amount of dust extinction with $A_V = 0.5$ mag to account for the non-negligible amount of extinction in quiescent galaxies observed at $z \geq 2$ (e.g. Whitaker et al. 2013; Marchesini et al. 2014). We adopt the dust extinction curve described by Calzetti et al. (2000).

Gas accretion rates on to the central galaxy are measured in spherical shells using the velocities and masses of gas particles. The bulk radial velocity of accreting or outflowing gas v is derived from the gas accretion rate $\dot{M}$, the width of the radial shell ΔR and the mass within the shell M as follows: $v = \Delta R\, \dot{M}/M$. Accretion rates and bulk velocities are corrected for the background Hubble flow.

3 COMPARISON WITH OBSERVATIONS

In this section, we demonstrate that many properties of the simulated massive galaxy evolve strongly over the redshift range $z = 2$–4. In addition, we show that by $z = 2$ stellar masses, sizes, and colours are broadly consistent with those of massive, quiescent galaxies observed at those redshifts.

The morphology of the simulated massive galaxy changes substantially in the $\sim$1.8 Gyr interval between $z = 4$ and 2, see Fig. 1. In particular, the galaxy evolves from a moderately massive ($M_* \sim 10^{10}$ M$_\odot$), compact ($R_{1/2} \sim 0.3$ kpc), blue ($U - V \sim 0.7$) galaxy with a disc component into a still compact ($R_{1/2} \sim 1$ kpc), massive ($M_* \sim 10^{11}$ M$_\odot$), early-type galaxy with significantly redder colours ($U - V \sim 1.25$). We summarize the evolution of many global properties of the simulated galaxy in Table 2.

Galaxies show a bimodal distribution of galaxy colours and SFRs both in the local Universe (e.g. Blanton et al. 2003; Kauffmann et al. 2003a; Balogh et al. 2004; Hogg et al. 2004) and out to at least $z \sim 2$ (e.g. Bell et al. 2004; Brammer et al. 2009, 2011; Taylor et al. 2009; Williams et al. 2009; Whitaker et al. 2011; Muzzin et al. 2013). Consequently, empirical cuts in colour–colour space or thresholds in sSFR are often used to classify galaxies as 'quiescent' or 'star forming'. A popular method of the first kind is a classification based on the $U - V$ and $V - J$ rest-frame colours, the so-called UVJ diagram (e.g. Labbé et al. 2005; Wuyts et al. 2007; Williams et al. 2009; van de Sande et al. 2013), which employs the fact that galaxy light dominated by an old stellar population shows a pronounced 4000 Å break in its spectrum. Cuts in sSFR are often based on a threshold of $\mathrm{sSFR}(z) = 0.3/t_\mathrm{H}$, where t_H is the age of the Universe at the given redshift (Franx et al. 2008; Fontana et al. 2009; Williams et al. 2010). This corresponds to $\sim 10^{-10}$ yr^{-1} at $z = 2$–3.

It should be kept in mind that the classification is based on the observed bimodality and not on matching galaxy properties to those of local star-forming or quiescent galaxies. In fact, galaxy properties change dramatically between $z \sim 2$ and today. For instance, star-forming galaxies at high-z have larger gas fraction (Daddi et al. 2008, 2010; Davé, Finlator & Oppenheimer 2012; Tacconi et al. 2010, 2013; Magdis et al. 2012; Feldmann 2013; Sargent et al. 2014), lower metallicities at a given stellar mass (Erb et al. 2006; Maiolino et al. 2008; Mannucci et al. 2010), and a more turbulent ISM (Förster Schreiber et al. 2006). High-redshift galaxies that are quiescent (based on the UVJ diagram) have younger stellar populations than local quiescent galaxies (e.g. Whitaker et al. 2013), significant dust extinction (e.g. Marchesini et al. 2014) and may also be forming stars at a substantial rate (e.g. Brammer et al. 2011; Straatman et al. 2014, but also cf. Utomo et al. 2014).

We explore the quiescent or star-forming nature of the simulated galaxy in the UVJ diagram of Fig. 2. The galaxy starts with blue $U - V$ and $V - J$ colours and, with time, turns redder in both colours. At $z \sim 2$ the galaxy crosses over into a region in colour–colour space that is associated with quiescent high-redshift galaxies. At this time its sSFR falls below $\sim 10^{-10}$ yr^{-1}, i.e. below the sSFR threshold mentioned above. Note, however, that the actual SFRs are still of the order of $\sim$10 M$_\odot$. Hence, although the galaxy qualifies as quiescent according to both the colour–colour and the sSFR criteria, it differs substantially from completely 'read and dead' galaxies in the local Universe. Reassuringly, the results for $z = 1$, see Fig. 2 and Table 2, show that the sSFR continues to decrease and the colours continue to redden with increasing cosmic time. We note that there is good agreement among the various runs.

A problem in many numerical simulations is that they tend to produce galaxies that are overly massive for their parent DM halo. This is a problem for Milky Way sized haloes (see Scannapieco et al. 2012, and references therein), for central galaxies in groups (e.g. Feldmann et al. 2010), but especially for cluster central galaxies (e.g. Borgani & Kravtsov 2011). Improved numerical resolution and modelling of stellar feedback has alleviated this problem

[1] More specifically, J refers to the NSFCam 'Mauna Kea' $J - MK$ filter, see http://irtfweb.ifa.hawaii.edu/~nsfcam2/Filter_Profiles.html.

[2] We chose the WFCam K-band filter used in the UKIRT Infrared Deep Sky Survey, see Hewett et al. (2006) and http://www.ukidss.org/technical/photom/photom.html.

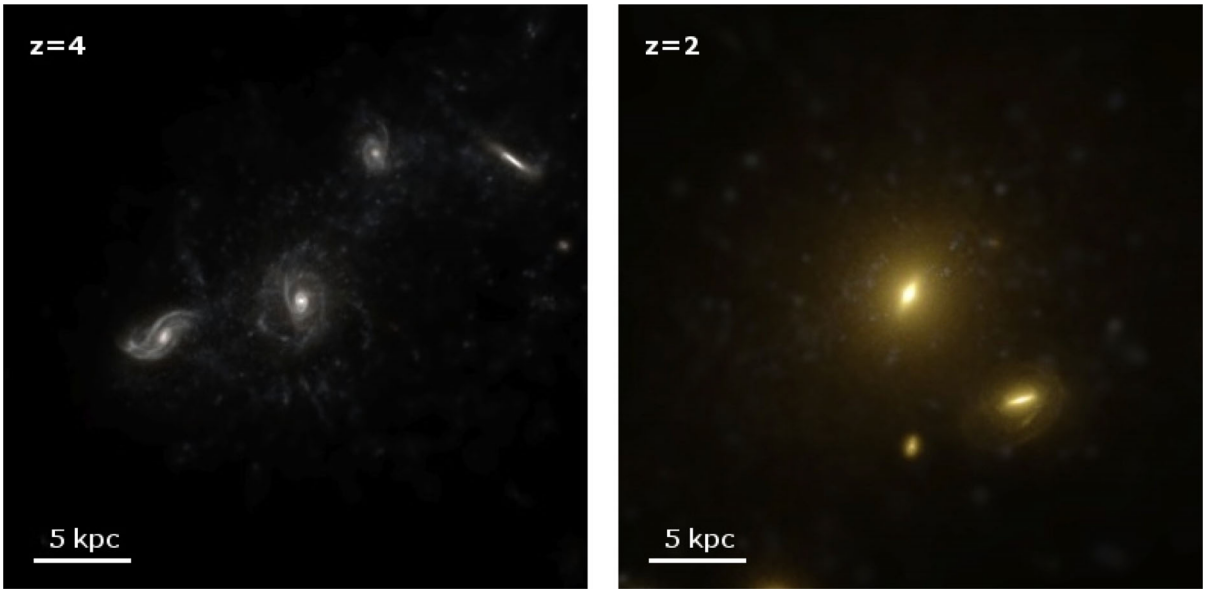

Figure 1. Evolution of the simulated massive galaxy (centre) over the redshift range $z = 4$ (left-hand panel, HR run) to $z = 2$ (right-hand panel, MR2 run). The panels show composite mock images in the *Hubble Space Telescope I*, *J* and *H* bands (logarithmic stretch). The size indicated at the bottom left is in proper kpc. Between $z = 4$ and 2 the massive galaxy evolves from a blue, compact, star-forming galaxy with a disc component into a much redder, still compact, early-type galaxy.

Table 2. Properties of the simulated central galaxy at selected redshifts. The first two columns label redshift and run, respectively. The remaining columns provide: (3) virial mass of the halo (the mean halo density is 180 times the background density of the Universe), (4) stellar mass within a 12 kpc radius (for $z \geq 2$) or a 20 kpc radius (for $z = 1$), (5) mass of gas colder than $T = 3 \times 10^4$ K within a 12 kpc radius ($z \geq 2$) or a 20 kpc radius ($z = 1$), (6) viral radius of the halo, (7) average projected radius containing half the stellar mass given in column (4), (8) sSFR within 12 kpc (for $z \geq 2$) or 20 kpc (for $z = 1$) averaged over the past 20 Myr, (9) same as column (8), but excluding star formation and stellar mass within the central 0.3 kpc, (10) WFCAM *K*-band magnitude in the observed frame within 12 kpc (for $z \geq 2$) or 20 kpc (for $z = 1$), (11–12) rest-frame Bessel *U*- and *V*-band magnitudes, and (13) rest-frame NSFCAM *J*-band magnitude. Magnitudes are given in the AB system and are derived based on the age, mass, and metallicity of each stellar particle using Bruzual & Charlot (2003) stellar synthesis models and a Chabrier (2003) IMF. We apply a uniform Calzetti et al. (2000) dust screen of $A_V = 0.5$ mag to account for a moderate amount of extinction observed in many quiescent high-redshift galaxies (Marchesini et al. 2014). We identify and mask any satellite galaxy before measuring the properties shown in columns 4 and 5 and in columns 7 through 13.

z	Run	log $M_{\rm vir}$ ($M_\odot$)	log M_* ($M_\odot$)	log $M_{\rm cgas}$ ($M_\odot$)	$R_{\rm vir}$ (kpc)	$R_{1/2}$ (kpc)	log sSFR (yr^{-1})	log sSFR' (yr^{-1})	$K_{\rm obs}$ mag_{AB}	$U_{\rm rest}$ mag_{AB}	$V_{\rm rest}$ mag_{AB}	$J_{\rm rest}$ mag_{AB}
5	LR	11.02	9.27	9.51	26.2	0.24	−8.05	−8.12	26.1	−20.4	−20.8	−20.8
5	MR1	10.98	9.33	9.49	25.4	0.26	−8.58	−8.32	26.5	−19.9	−20.5	−20.6
5	MR2	11.03	9.47	9.34	26.4	0.20	−8.54	−8.84	26.3	−20.1	−20.7	−20.9
5	HR	10.97	9.31	9.45	25.1	0.21	−8.44	−8.65	26.6	−19.9	−20.4	−20.6
4	LR	11.83	10.17	9.87	59.0	0.28	−8.77	−8.84	24.1	−21.4	−22.1	−22.4
4	MR1	11.82	10.19	9.77	58.6	0.36	−8.83	−8.94	24.0	−21.4	−22.2	−22.4
4	MR2	11.82	10.27	9.86	58.7	0.26	−8.84	−8.79	23.9	−21.5	−22.3	−22.6
4	HR	11.82	10.00	9.82	58.4	0.33	−8.59	−8.47	24.4	−21.1	−21.7	−21.9
3.5	LR	12.03	10.73	9.83	76.4	0.51	−8.66	−8.85	22.3	−22.9	−23.6	−23.9
3.5	MR1	12.03	10.65	9.76	76.5	0.55	−8.58	−9.01	22.5	−22.8	−23.4	−23.7
3.5	MR2	12.03	10.72	9.85	76.5	0.39	−8.76	−8.96	22.4	−22.7	−23.4	−23.7
3.5	HR	12.03	10.59	9.67	76.3	1.00	−8.74	−8.87	22.8	−22.4	−23.1	−23.4
3.4	LR	12.05	10.76	9.74	79.4	0.43	−8.98	−9.49	22.4	−22.7	−23.5	−23.8
3.4	MR1	12.05	10.67	9.52	79.5	0.51	−8.99	−9.43	22.6	−22.4	−23.2	−23.5
3.4	MR2	12.05	10.74	9.67	79.5	0.34	−9.16	−9.38	22.5	−22.5	−23.3	−23.7
3.4	HR	12.05	10.59	9.52	79.1	0.83	−9.13	−9.22	22.9	−22.1	−22.9	−23.2
3	LR	12.10	10.82	9.55	90.1	0.43	−9.26	−9.85	22.3	−22.3	−23.3	−23.7
3	MR1	12.12	10.71	9.34	91.6	0.56	−9.38	−9.85	22.7	−21.9	−22.9	−23.3
3	MR2	12.12	10.78	9.61	91.5	0.44	−9.37	−9.67	22.5	−22.0	−23.1	−23.5
2	LR	12.45	11.01	9.64	154.0	0.96	−9.65	−9.78	21.5	−21.8	−23.0	−23.7
2	MR1	12.47	10.89	9.33	156.8	0.84	−9.73	−9.86	21.8	−21.4	−22.7	−23.4
2	MR2	12.47	10.96	9.32	156.2	0.77	−9.66	−10.27	21.6	−21.7	−22.9	−23.6
1	LR	12.86	11.29	9.39	323.7	1.71	−10.22	−10.60	19.6	−21.4	−22.9	−23.8

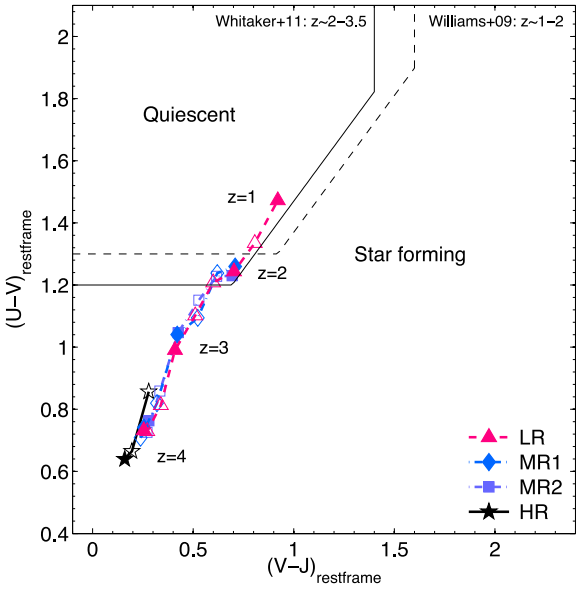

Figure 2. Rest-frame *UVJ* colours of the simulated galaxy from $z = 4$ to 1. Symbols connected with solid, dot–dashed, and dashed lines show the $U - V$ and $V - J$ colours of the central galaxy in the HR, MR, and LR runs (see legend). The filled symbols mark the specific redshifts 4, 3, 2, and 1 (from bottom left to top right). Observed galaxies are often classified as quiescent or star forming based on their position in the *UVJ* diagram. We show two different selection choices: one for galaxies at $z \sim 2$–3.5 (solid black line; Whitaker et al. 2011) and one for galaxies at $z \sim 1$–2 (dashed black line; Williams et al. 2009), see legend. At $z \lesssim 2$, the simulated galaxy would likely be classified as a quiescent galaxy based on its location in the *UVJ* diagram.

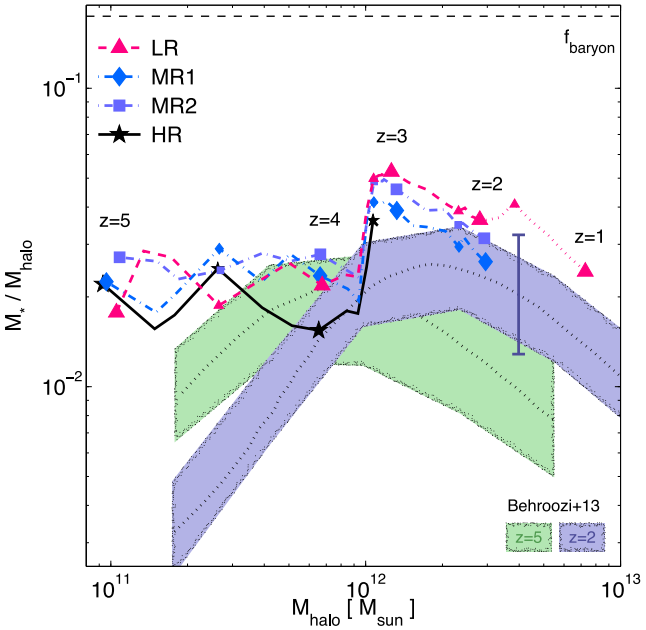

Figure 3. The stellar-to-virial mass ratio as function of virial mass. Symbols connected with solid, dot–dashed, and dashed lines show the stellar-to-virial mass ratio of the central galaxy in the HR, MR, and LR runs (see legend). The symbols mark the specific redshifts 5, 4, 3, 2, and 1 (from left to right). The thin, horizontal dashed line at the top shows the universal baryon fraction. The dotted lines and shaded regions are the mean value of the stellar fraction and its combined systematic and statistical uncertainty range, respectively, as derived from abundance matching techniques (Behroozi et al. 2013). The error bar on the right shows the typical scatter (0.2 dex) for individual haloes (More et al. 2009; Reddick et al. 2013) as comparison. The stellar mass fraction of the simulated galaxy evolves only weakly and is in agreement with abundance matching predictions over the redshift range $z = 4$ to 2.

substantially for galaxies in Milky Way sized (and smaller) haloes (Guedes et al. 2011; Brook et al. 2012; Aumer et al. 2013; Hopkins et al. 2014; Munshi et al. 2013) and AGN feedback offers a potential solution for galaxies in clusters (e.g. Teyssier et al. 2011; Martizzi, Teyssier & Moore 2012). More specifically, an important success of the Guedes et al. (2011) model is that the obtained stellar mass at $z = 0$ is consistent with the observed stellar mass of the Milky Way. However, a priori it is unclear, whether this model, which includes thermal feedback from SNe as its only energetic feedback source, is adequate to prevent overcooling in group sized haloes, even at high redshift (see Stinson et al. 2013).

Hence, we compare in Fig. 3 the stellar-to-virial mass ratio of the simulated galaxy in the HR, MR, and LR runs with the relation derived by Behroozi, Wechsler & Conroy (2013) using the abundance matching technique (Vale & Ostriker 2004; Conroy, Wechsler & Kravtsov 2006; Behroozi, Conroy & Wechsler 2010; Guo et al. 2010; Moster et al. 2010; Trujillo-Gomez et al. 2011; Moster, Naab & White 2013; Reddick et al. 2013). The HR run of the Argo simulation features the same physics and has the same resolution as the simulation by Guedes et al. (2011), and thus is a solid baseline against which we can cross-check our lower resolution runs.

As Fig. 3 shows, the stellar-to-virial mass ratio of the simulated central galaxy is in good agreement with the empirically derived estimates by Behroozi et al. (2013) over the $z = 4$ to 1 range. At higher redshifts the stellar-to-virial mass ratio in the simulation lies above abundance matching predictions. Interestingly, the stellar-to-virial mass ratio is almost constant with a value of ∼2–3 per cent, although it shows fluctuations of the order of the scatter in the empirically derived stellar-to-virial mass relation. The rapid increase in the stellar mass at $z \sim 3.5$ results from a massive starburst and several galaxy mergers that add stellar mass to the central galaxy but do not change the virial mass significantly. Later on, the increase in stellar mass of the galaxy is outpaced by the increase in virial mass and the stellar-to-virial mass ratio drops.

The virial mass among the different runs is in excellent agreement. The stellar mass is also agreeing well, although there is a visible trend of increased stellar mass with decreased numerical resolution, at least after $z = 4$. For instance, at $z = 3.5$ the stellar mass is ∼40 per cent larger in the LR run than in the HR run. Overall, the agreement is reasonable given that gas particle masses change by a factor of 64 between the LR and the HR run. We note that our LR run has a significantly better resolution than numerical studies of large ensembles of galaxies in representative cosmological volumes (e.g. Haas et al. 2013; Le Brun et al. 2014; Vogelsberger et al. 2013). Hence, simulations that use a physical model similar to ours, but at much lower resolution, may run the risk of overestimating the importance of additional feedback mechanisms to reduce stellar masses of galaxies to observed levels. For simulations that reach the resolution of our MR or HR runs, a simple SN-based feedback model (Stinson et al. 2006; Guedes et al. 2011) may be all that is needed to reproduce adequate stellar masses for galaxies residing in $\sim 10^{12.5}$ $M_\odot$ haloes at $z = 2$.

Another interesting property of massive quiescent galaxies at redshift $z \sim 1$–3 is that they are significantly more compact than passive galaxies of similar mass at $z \sim 0$ (e.g. Daddi et al. 2005; Trujillo et al. 2006; Toft et al. 2007; Cimatti et al. 2008; Franx et al. 2008; van der Wel et al. 2008; van Dokkum et al. 2008, 2010;

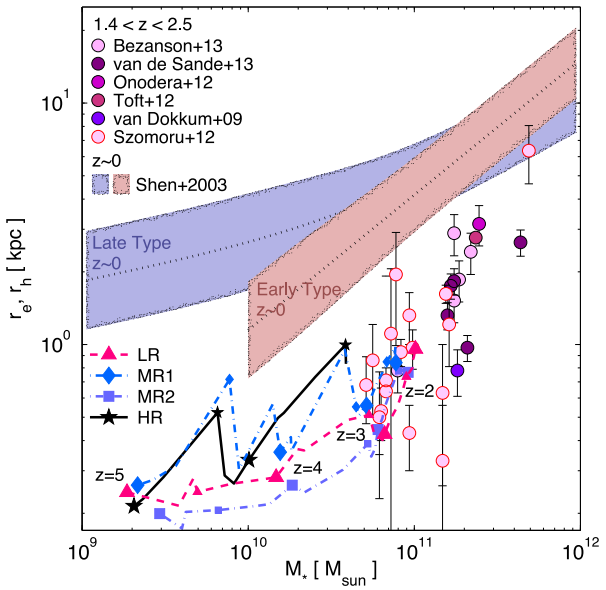

Figure 4. Stellar mass–size relation of simulated and observed galaxies. Symbols connected with solid, dot–dashed, and dashed lines show the half-mass radius of the central galaxy in the HR, MR, and LR runs (see legend). The symbols mark the specific redshifts 5, 4, 3, and 2 (from left to right). The circles with error bars show the effective radii and stellar masses of $z \sim 2$ quiescent galaxies (van Dokkum et al. 2009; Onodera et al. 2012; Szomoru et al. 2012; Toft et al. 2012; Bezanson et al. 2013; van de Sande et al. 2013). The dotted lines and shaded regions show the observed $z \sim 0$ mass–size relation and its scatter (Shen et al. 2003) for galaxies with Sérsic index $\geq$2.5 (early types) and <2.5 (late types). The size of the simulated galaxy at $z \sim 2$–3 is in good agreement with observations of quiescent, similarly massive galaxies at $z \sim 2$. In particular, it is more compact (by 0.3–0.5 dex) than early-type galaxies of similar mass in the local Universe.

Cassata et al. 2011; Damjanov et al. 2011; Szomoru, Franx & van Dokkum 2012; Carollo et al. 2013). This discovery, first based on measuring surface brightness profiles, was subsequently confirmed spectroscopically (e.g. van Dokkum, Kriek & Franx 2009; Newman et al. 2010; van de Sande et al. 2011; Onodera et al. 2012; Toft et al. 2012; Bezanson et al. 2013; van de Sande et al. 2013).

In Fig. 4 we test how well the size of our simulated galaxy fares with those of massive, quiescent galaxies at high $z \sim 2$. At $z = 2$ our simulated galaxy has size of $\sim$1 kpc and a stellar mass of $\sim 10^{11}$ $M_\odot$. This agrees well with the sizes of observed, similarly massive, quiescent galaxies. However, we caution the reader that our sizes are half-mass radii and not 'effective radii'. The latter are typically used in the observational literature and are derived from a parametric fit to the light surface profiles. While keeping these caveats in mind, the similarity of the sizes at $z \sim 2$ is encouraging.

At $z > 3.5$ the galaxy is already rather compact but (as we show later) its SFR and stellar mass place it on the star-forming sequence, i.e. at $z \sim 4$ it is a compact star-forming galaxy. Such galaxies, sometimes called blue nuggets, are possibly an important intermediate stage preceding quenching and the formation of compact, passive galaxies (Wuyts et al. 2011; Barro et al. 2013; Patel et al. 2013).

The size of our galaxy at $z \sim 2$–3 changes little among the runs, again showing that we have reasonably well converged to a final answer by that time. At higher redshifts, however, there are noticeable differences of up to a factor of 3 between the runs. Note, that the half-mass radius is barely resolved at $z > 4$ even in our HR run. Overall, we find a tendency for larger size with increasing resolution. The trend is possibly caused by the enhanced efficiency of SN-driven winds to remove low angular momentum material with increased numerical resolution, as seen in previous work (e.g. Governato et al. 2010; Guedes et al. 2011). In addition, a higher numerical resolution results in a larger number of resolved satellite galaxies and, hence, may drive larger sizes via an increase in the minor merging rate.

The size evolution in our HR run and in the MR1 run shows some interesting trends. First, the half-mass radius grows between $z = 5$ and 4.5 as a consequence of star formation in the outer region of the galaxy. Then at $z = 4.5$ the galaxy undergoes a significant stellar merger and the half-mass radius drops significantly. This behaviour repeats itself, when the galaxy size increases between $z \sim 4.3$ and 3.4, only to start to drop at $z \sim 3.5$, when (as we show later) the galaxy undergoes a critical event in its history. Between $z = 3.5$ and 3 the evolution in half-mass radius and stellar mass is minimal. After $z = 3$ the galaxy grows quickly in size, but adds little stellar mass. If we trust the size evolution in the HR and MR1 runs at $z \geq 3.5$, then this points to a picture in which galaxies grow in size while being on the star-forming sequence (via star formation), shrink when engaging in wet mergers, and then finally grow via gas-poor mergers after leaving the star-forming sequence.

So far, we have tested how the stellar mass and the structural properties of the simulated massive galaxy compare with observational data. Another crucial check is to test the properties of the hot gas surrounding the galaxy. The entropy profile is of particular importance, because, it measures the cooling time of the circumgalactic gas.

In Fig. 5 we show the entropy profile at various redshifts and compare it with the entropy profile of the nearby massive, quiescent elliptical galaxy NGC 720 (Humphrey et al. 2011). NGC 720 has a stellar mass of $\sim 10^{11}$ $M_\odot$ and is therefore a potential low-redshift counterpart of our simulated galaxy at $z = 2$. As the figure demonstrates, the entropy profile of the simulated galaxy at $z = 2$ and the entropy profile of NGC 720 have the same shape and almost the same normalization (to within a factor of 2). The intriguing property of NGC 720 is that it somehow prevents the overcooling problem without increasing the entropy of the gas to the critical value of $\sim$100 KeV cm^2, above which cooling would take longer than a Hubble time. Hence, it is a counter example to the picture in which star formation is quenched because of a previous quasar mode heating phase (e.g. Scannapieco & Oh 2004).

In fact, the cooling times of the gas are quite short, especially at small distances from the centre of the galaxy. At $z = 2$, the cooling time is only $\sim$10 Myr for $r = 1$ kpc, while it reaches several Gyr ($\sim$a Hubble time) at $r = 100$ kpc. This indicates that to avoid overcooling, gravitational heating as well as heating from SNe have to replenish a large fraction of the cooling losses. We note that the cooling time generally exceeds the dynamical time by a factor of a few at each radius. The heating mechanisms implemented in the Argo simulation might not suffice in preventing overcooling at late times. In the real Universe, additional heating sources, e.g. an AGN in radio mode, may step in and supplant the decreasing contribution from gravitational heating (fewer mergers) and SN heating (fewer core collapse SNe).

The entropy profile of the MR and HR runs are converged except maybe in the central kpc. This difference, however, may explain the somewhat larger stellar mass in the MR2 run compared with the HR run as most of the star formation takes place in the central kpc. We also note that the entropy is lower in the LR run (by a factor of $\sim$2–3 at most radii) although it matches the entropy of the MR2 run in the central region. We interpret the smaller entropy in the LR and MR2 runs as a consequence of the reduced efficiency of

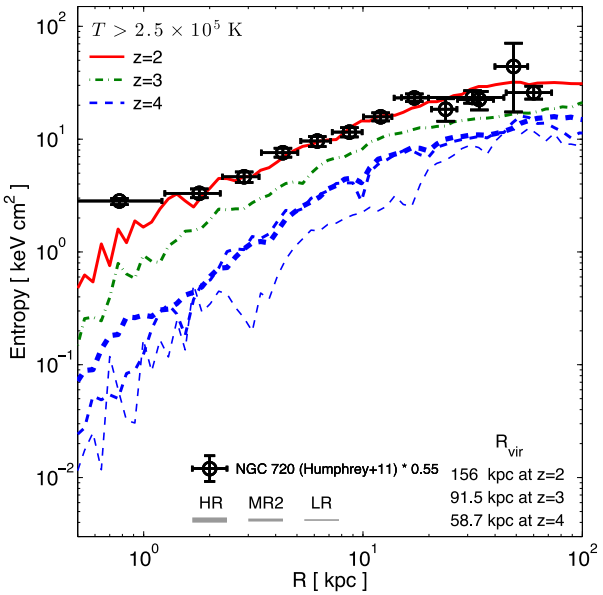

Figure 5. Entropy profile ($S = T/n_e^{2/3}$, where n_e is the number density of electrons) radially averaged in spherical shells. The red solid, green dot–dashed and blue dashed lines show the entropy profile of hot gas ($T > 2.5 \times 10^5$ K) around the simulated massive galaxy at redshift 2, 3, and 4, respectively. At $z = 4$ the thickness of the line corresponds to the resolution of the simulation, see legend. The lines at $z = 2$ and 3 are from the MR2 run. The circles with error bars show the entropy profile (re-scaled to 55 per cent to match the normalization) of the hot gas around the nearby quiescent elliptical galaxy NGC 720 (Humphrey et al. 2011). The shape of the entropy profiles of the simulated galaxy at $z = 2$ and the observed galaxy at $z \sim 0$ agree well, except maybe in the central kpc of the galaxy. The entropy is below the critical value of $\sim$100 keV cm^2 (Voit & Bryan 2001; Scannapieco & Oh 2004) indicating that the gas cooling time is shorter than the Hubble time. SN feedback as well as gravitational heating balance the cooling rate of the hot gas halo and avoid excessive net cooling rates at $z \sim 2$. At low redshifts additional feedback sources, e.g. a radio mode AGN, are likely required to maintain the low SFR of the central galaxy.

SN feedback at lower resolution and with a lower star formation threshold (e.g. Guedes et al. 2011; Munshi et al. 2013).

4 ON LEAVING THE STAR-FORMING SEQUENCE

In the previous section, we traced the simulated massive galaxy across $z = 5$–2 and analysed how its properties change over this redshift range, finding significant evolution. We showed that stellar mass, stellar-to-virial mass ratio, size, and colours of the simulated galaxy are in good agreement with those of massive ($M_* \sim 10^{11}$ M$_\odot$) quiescent galaxies observed at high redshift. We will now discuss the evolution of the SFR and the sSFR of the simulated galaxy.

We show in Fig. 6 the sSFR of the simulated galaxy as function of its stellar mass. At $z \geq 3.5$ the galaxy lies on the observed star-forming sequence (sometimes called the 'main sequence') with an sSFR of $\sim 2 \times 10^{-9}$ yr^{-1}. We should point out that this is an important achievement of the Argo simulation. Indeed, published galaxy simulation studies, especially those with resolutions worse than our LR run, often obtain sSFRs that lie off the observed relation by a factor of 2–3 (e.g. Finlator, Oppenheimer & Davé 2011; Haas et al. 2013).

At $z = 3.5$ the galaxy leaves, rather abruptly, the star-forming sequence. The sSFR drops by a large factor ($\sim$5–10) within a few hundred Myr. At $z = 2$, the sSFR is only $\sim 10^{-10}$ yr^{-1}, i.e. more than an order of magnitude lower than the sSFRs of galaxies on the star-forming sequence. The abrupt suppression of the sSFR at $z \sim 3.5$–3 is reproduced in all our runs. The drop is somewhat steeper in the MR and HR runs compared with the LR run. We interpret this result as a consequence of SN feedback being somewhat more efficient in our higher resolution runs, see also Section 3.

However, *the drop of the sSFR is not triggered by SN feedback*. This is shown in the right-hand panel of Fig. 6 where we test how the sSFR evolves after turning off the feedback at $z = 4$. We find little difference whether SN feedback is active or not.

Feedback nonetheless plays two important roles. First, feedback at early times ($z \gg 4$) is required to prevent overcooling and to reproduce sSFRs consistent with observations. We illustrate this point in Fig. 6 where we show re-simulations of the LR run without feedback. The run in which we switch off SN feedback after $z = 4$ differs little from the default LR run as noted above. In contrast, the run without any SN feedback (LR-noFB) results in too high a stellar mass which outweighs also the somewhat higher SFR, especially at early times. As a consequence, the sSFR at $z > 3.5$ is significantly reduced in this run.

Fig. 6 shows that the sSFR levels off at $\sim 10^{-10}$ yr^{-1} (at $z = 2$) and only slowly decreases as time proceeds. As we will demonstrate in Section 5, this floor in the sSFR at high redshift is sustained by the accretion of cool gas from large distances. However, one may wonder whether mass-loss from evolved stars (included in our models) contributes at a significant level (e.g. Fumagalli et al. 2014). The right-hand panel of Fig. 6 shows a re-simulation of the LR run, but with the stellar mass-loss switched off after $z = 3$. In this case the sSFR levels off at a somewhat smaller value compared with the default LR run. However, the difference is small (smaller than a factor of 2) proving that the $\sim 10^{-10}$ yr^{-1} floor of the sSFR at $z = 2$ is not related to a replenishment of the ISM by mass-loss from an evolved stellar population.

The left-hand panel of Fig. 7 shows the star formation history of the massive galaxy traced through time. The SFR increases steeply from $\sim$10 M$_\odot$ yr^{-1} at $z = 5$ to $\sim$100 M$_\odot$ yr^{-1} at $z = 3.5$. Afterwards the SFR decreases and settles at $\sim$15 M$_\odot$ yr^{-1} at $z = 2$ (for the MR runs).

Comparing the LR, MR and HR runs, we find differences in the SFR (close to a factor of 2), especially around the peak of SFR at $z \sim 3.5$ and in the subsequent star formation activity at $z \sim 2$–3.5. The general trend is that the SFR is reduced in the runs with higher spatial resolution, i.e. in the runs with more efficient stellar feedback. However, a significant fraction of the star formation takes place in the central softening length of the simulation and, hence, is unresolved. We thus show in the right-hand panel of Fig. 7 the SFR that occurs outside the central 300 pc. In this case, the predictions of the various runs are well converged.

The SFR can be approximated by an exponential with constant offset, i.e.

$$\mathrm{SFR}(t) = \mathrm{SFR}(t_*) \left[(1 - c) e^{\frac{t - t_*}{\tau}} + c \right]. \quad (1)$$

Here, $t_* = t(z = 3.5) = 1.9$ Gyr, c is an offset, and τ is the star formation time-scale. The value of τ is positive for an exponentially growing SFR and negative for an exponentially declining SFR. The absolute value of τ measures the gas accretion time ($M_{\rm gas}/\dot{M}_{\rm gas,acc}$) and the gas depletion time ($M_{\rm gas}$/SFR) in the growing and declining phase, respectively, see Appendix A.

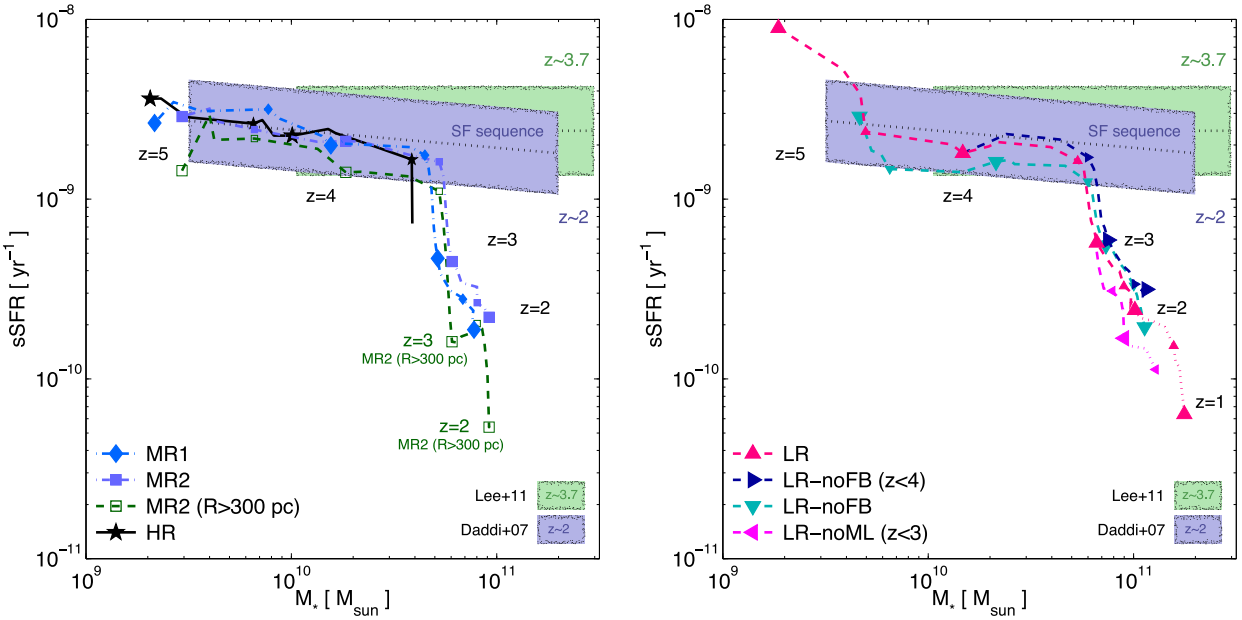

Figure 6. Evolution of the sSFR. Left-hand panel: evolution of the sSFR of the central galaxy in the HR (black solid line, star symbols) and the MR (dot–dashed lines, square and diamond symbols) runs. The line connecting the empty squares show the sSFR of the central galaxy when the star formation and stellar mass within the central 300 pc are excluded. Right-hand panel: evolution of the sSFR of the central galaxy in the LR run (red lines, upward-pointing triangle symbols) and in various re-simulations of the LR run. The blue line with rightward pointing triangles is a run that starts with the $z = 4$ snapshot of the LR run and the continues down to $z = 2$, but with SN feedback turned off. Similarly, the green line with downward-pointing triangles is a re-simulation of the LR run but without SN feedback from the start. The magenta line with leftward pointing triangles is a re-simulation that starts with the $z = 3$ snapshot of the LR run but with stellar mass-loss turned off. In both panels, large symbols mark redshifts 5, 4, 3, 2 and 1 (from left to right), small symbols mark redshifts 4.5, 3.5, 2.5, 1.5. The SFRs are measured within a 20 Myr time interval and a five-point moving average is applied to the sSFR to reduce short-term fluctuations and highlight the overall evolutionary trend. The dotted lines and shaded regions show the mean location and the 1σ scatter, respectively, of the observed star formation sequence at $z \sim 3.7$ (Lee et al. 2011) and $z \sim 2$ (Daddi et al. 2007). Stellar masses are converted to a Chabrier (2003) IMF. The central galaxy evolves along the observed star-forming sequence until $z \sim 3.5$. At $z \sim 3.5$ the sSFR drops by a large factor (∼5–10) within a few 100 Myr, resulting in a massive galaxy with strongly suppressed star formation activity.

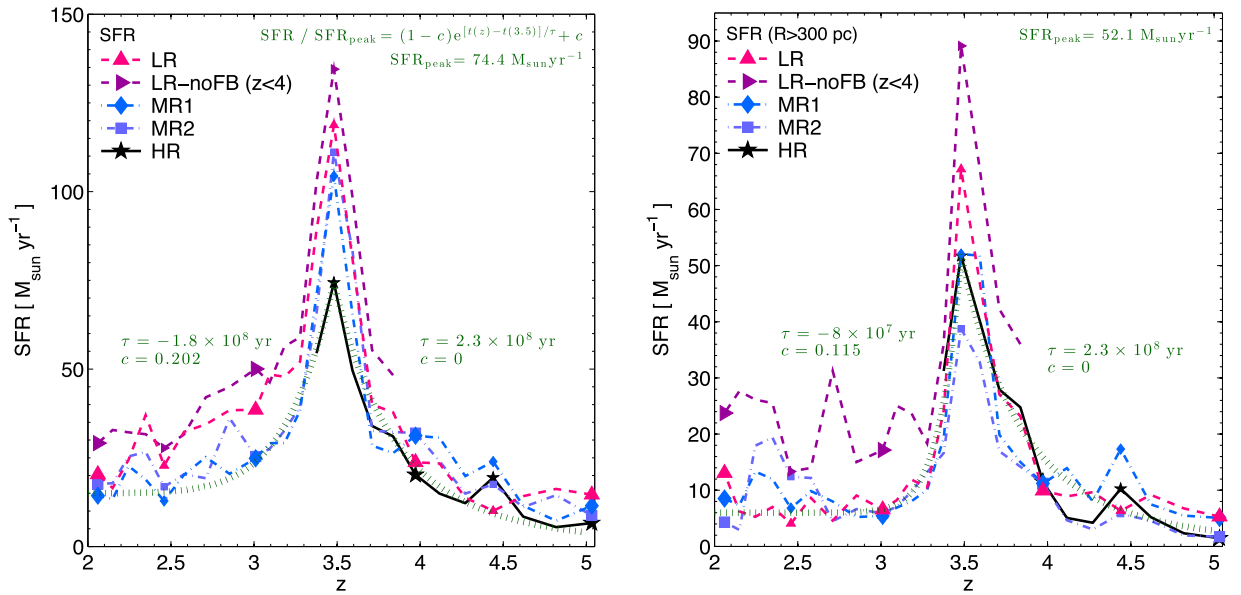

Figure 7. Star formation history of the simulated massive galaxy. Left-hand panel: SFR within a radius of 12 kpc excluding contributions from satellite galaxies. Right-hand panel: SFR within a shell with a 0.3 kpc inner radius and a 12 kpc outer radius. Symbols connected with the solid, dot–dashed, and dashed lines show the SFR of the central galaxy in the HR, MR, and LR runs (see legend). The SFR increases exponentially until $z = 3.5$ after which it declines. At $z \sim 2$–3 a non-negligible fraction of the star formation takes place within the central 0.3 kpc.

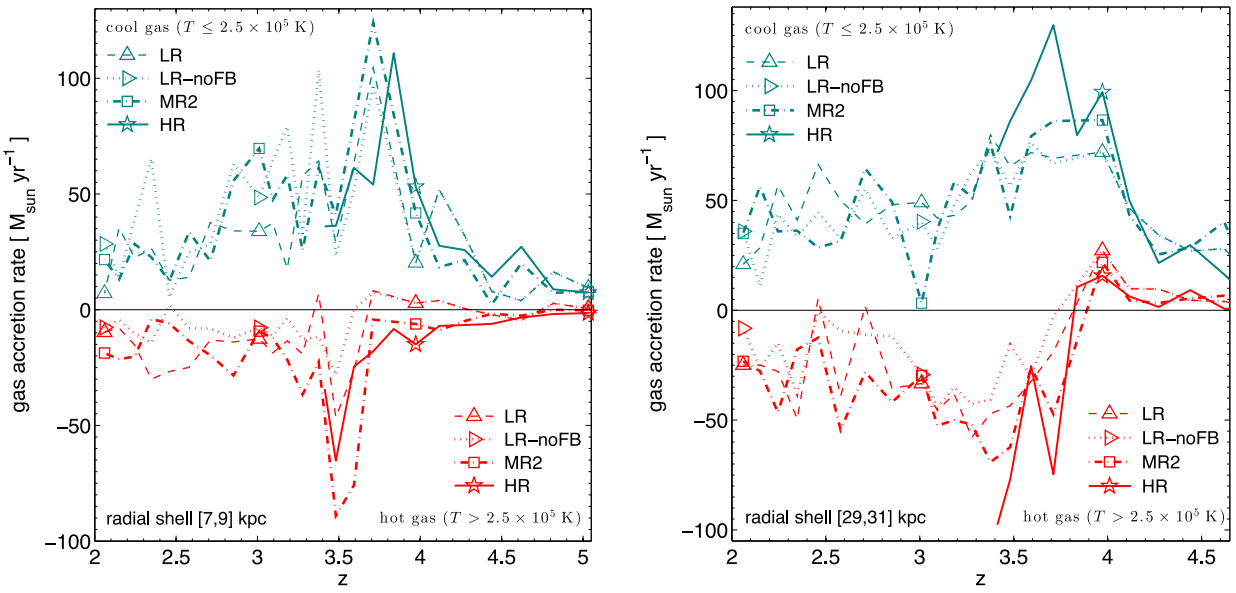

Figure 8. Accretion rates of cool ($T \leq 2.5 \times 10^5$ K) and hot ($T > 2.5 \times 10^5$ K) gas as measured in spherical shells centred at proper 8 kpc (left-hand panel) and 30 kpc (right-hand panel) for the LR, MR, and HR runs (see legend). Cool gas flows towards the central galaxy, while hot gas, especially at $z \lesssim 4$, preferentially flows outward. At $z \sim 2$–3 the outflow rate of hot gas balances a large fraction of the inflow rate of the cool gas at each radius, resulting in little net accretion of gas towards the galaxy.

At $z > 3.5$ the SFR increases roughly exponentially with time ($c = 0$). The e-fold time is $\tau \sim 2.3 \times 10^8$ yr for the HR run. This e-fold time corresponds to an sSFR of $1/\tau \sim 4 \times 10^{-9}$ yr^{-1}, approximately a factor of 2 larger than the actual sSFR of the galaxy, see Fig. 6. However, the $1/\tau$ estimate ignores the contribution of stellar mass growth via mergers.[3] Indeed, the simulated galaxy experiences a number of significant merger events between $z = 5$ and 3.5. It undergoes a major galaxy merger at $z = 4.4$ (the stellar mass ratio is 1: 2.5). It is involved further in a 1: 10 merger at $z = 3.9$, a 1: 7 merger at $z = 3.7$, and a 1: 5 merger at $z = 3.5$. The overall stellar mass growth provided by galaxy mergers is $\sim$1.7–2 $\times$ 10^{10} $M_\odot$, comparable to the mass growth by *in situ* star formation ($\sim$1.6 $\times$ 10^{10} $M_\odot$). The mass acquisition via mergers explains why the sSFR is actually a factor of 2 smaller than the estimate $1/\tau$.

The exponential growth of the SFR at $z > 3.5$ also ensures that the sSFR remains roughly constant from $z = 5$ to 3.5. Naively, one might expect that a peak in the SFR translates into a peak in the sSFR, i.e. in an upward deviation from the star-forming sequence at $z \sim 3.5$–4. However, since the exponential function is proportional to its own derivative the sSFR is constant for a galaxy that grows most of its mass via an exponentially increasing SFR.

The star formation activity peaks at $z = 3.5$. In the HR run, the SFR reaches $\sim$75 $M_\odot$ yr^{-1} ($\sim$52 $M_\odot$ yr^{-1} if we exclude the central 300 pc). Afterwards the SFR declines rapidly (approximately as an exponential with constant offset). The SFR has an e-fold time of $\sim$200 Myr (100 Myr if we exclude non-central star formation). Hence, within a few 100 Myr the SFR decreases to the much lower levels of $\sim$15–20 $M_\odot$ yr^{-1} ($\sim$5–10 $M_\odot$ yr^{-1} if we exclude non-central star formation).

Although the galaxy leaves the star-forming sequence at $z \sim 3.5$ (Fig. 6), and reduces its SFR within a few 100 Myr (Fig. 7), it would not be identified as a quiescent galaxy based on the *UVJ* classification (Fig. 2) until $z \sim 2$. We suggest that this classification be improved by extending the region of quiescent galaxies in the *UVJ* diagram to the lower left.

[3] It also ignores stellar mass-loss, which, when properly taken into account, increases the discrepancy by a few tens of a per cent.

5 THE ORIGIN OF STAR FORMATION QUENCHING

The sSFR of the simulated massive galaxy decreases steeply at $z \sim 3.5$. We showed in the previous section that this decline is not triggered by feedback processes. Instead, we will now demonstrate that the reduction of the gas supply from the cosmic web, a process we dub 'cosmological starvation' (or 'cosmic starvation'), is to blame.

In Fig. 8 we plot the accretion rates of cool ($T \leq 2.5 \times 10^5$ K) and hot ($T > 2.5 \times 10^5$ K) gas through spherical shells centred on the galaxy. The cool gas accretion rate clearly mimics the star formation history, both in shape and normalization, especially if the accretion rate is measured close to the galaxy (here at 8 kpc distance). The delay ($\sim$100 Myr) between the peak in the cool gas accretion rate ($z \sim 3.7$–3.8) and the peak in the SFR ($z \sim 3.5$) indicates the time it takes the cool gas to reach the galaxy and participate in star formation.

Hot gas flows towards galaxy only at $z > 4$ and only if it is located sufficiently far from the massive galaxy. At $z < 4$ and/or at small radii, the bulk motion of hot gas is to move away from the galaxy. The outflow rate at small radii tracks the SFR rather precisely with a mass loading factor of the order of unity. The outflow rate peaks at $z \sim 3.5$ with values of 50–100 $M_\odot$ yr^{-1}. At larger radii the outflow rate evolves more gradual with time. At late times the outflow rate of the hot gas balances a substantial fraction of the cool gas accretion rate. Hence, the net gas accretion at late times is rather small (typically no more than 15 $M_\odot$ yr^{-1}).

The cool gas accretion rates agree well among the various runs. The outflow rates of the hot gas differ noticeably, however. In particular, the LR run shows weaker hot gas outflows compared with the MR and HR runs. Again, this is consistent with our overall

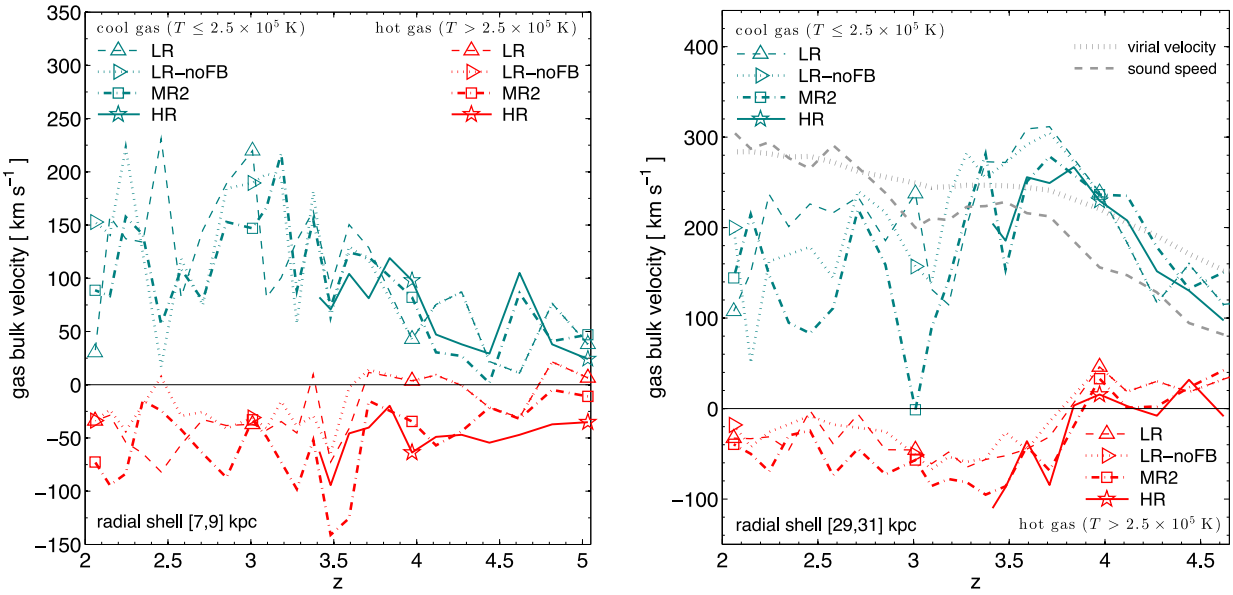

Figure 9. Bulk radial motion of cool ($T \leq 2.5 \times 10^5$ K) and hot ($T > 2.5 \times 10^5$ K) gas measured in spherical shells centred at proper 8 kpc (left-hand panel) and 30 kpc (right-hand panel) for the LR, MR, and HR runs (see legend). The thick dashed lines show the virial velocity ($[GM_{\rm halo}/R_{\rm halo}]^{1/2}$; upper line) and the adiabatic sound speed ($[\frac{5}{3}k_{\rm B}\overline{T}/\mu]^{1/2}$, with $\mu = 0.6m_{\rm H}$ and $\overline{T}$ the mass-weighted average temperature in the shell; lower line), respectively. Bulk velocities are measured in 2 kpc thick spherical shells. At $z > 3.5$ cool gas at large distances flows inward supersonically w.r.t. the hot halo surrounding the galaxy. Near the centre of the halo and at late times the inflow rate is subsonic. At $z < 3.8$ and/or at small radii hot gas is moving radially outward with an average bulk velocity of ∼50–100 km s^{-1} (in the MR and HR runs).

finding that the SN feedback scheme we use in the Argo simulation is less effective at LR. Nonetheless, there is a bulk radial motion of the hot gas even if we turn off the feedback completely. We suspect that some fraction of the hot gas bulk motion is related to the push of the thermal pressure of hot gas and to sloshing motions of the gas induced by infalling substructures. SN feedback, however, is responsible for the much stronger hot gas outflows in the MR and HR runs. We note that the hot gas outflows do not appear to strongly affect the star formation activity of the massive galaxy at $z = 3.3$–5, when much of the stellar mass of the massive galaxy is built. At late times ($z = 2$–3.3) the larger outflow rates in the MR2 run compared with the LR run result in a reduced star formation activity, but the overall impact on the $z = 2$ stellar mass is relatively small, see Table 2.

To understand the nature of the cool gas inflows and the hot gas outflows we plot in Fig. 9 their bulk velocities. The radial velocity of the cool gas increases until $z \sim 3.5$–3.8. By $z \sim 3.5$ the inflow velocity is ∼250–300 km s^{-1} at a 30 kpc distance and ∼100–150 km s^{-1} at a 8 kpc distance from the centre of the halo, respectively. The bulk flow of the hot gas is directed away from the massive galaxy, with flow velocities of ∼50–100 km s^{-1}. These velocities are relatively small indicating that SN feedback in our simulation probably primarily heats the hot gas halo and induces a pressure supported expansion of the hot halo atmosphere surrounding the central galaxy.

The inflow velocity of the cool gas stalls at late times. Combined with the sharp decrease in the net mass accretion rate this indicates that the mean density of the gas accreted at late times is lower. A lower density allows the virial shock to grow to larger radii (Birnboim & Dekel 2003, equation 29). The virial shock heats a much larger volume of gas surrounding the galaxy to the virial temperature of a few 10^6 K, significantly reducing the cooling rate of gas at large distances from the galaxy. The virial shock thus helps in reducing the inflow rate of cool gas from large to small radii. However, ultimately these events are set in place by the reduction of the gas accretion rates at large radii and not by the formation of the virial shock. In fact, a virial shock forms well before $z = 4$ when the simulated galaxy is forming stars at an increasing rate and its sSFR matches those of observed galaxies on the star-forming sequence.

Furthermore, the reduction of the gas accretion and the subsequent drop in the SFR is not merely a consequence of hydrodynamics and gas cooling, but related to gravitational processes, as we show now. Fig. 10 plots the evolution of the virial mass of the halo as well as the evolution of the mass of DM, baryons (gas and stellar matter), and cool gas enclosed in *fixed proper radii*.

The DM mass, the baryonic mass, the cool gas mass, and the virial mass all grow at $z \gtrsim 3.7$. The DM mass and the cool gas mass increase roughly in look-step, while the baryonic mass and the virial mass grow somewhat faster. Hence, at $z \gtrsim 3.7$ the halo is in a phase of fast accretion (*'collapse phase'*). Note that the mass increases within any fixed proper radius $r < R_{\rm vir}$ indicating genuine growth on all scales. During the collapse phase, the gas accretion rates and the SFR increase approximately exponentially and much of the *in situ* formed stellar mass of the galaxy is built.

At $z \lesssim 3.5$ this phase of fast accretion comes to a halt. This can be seen clearly in the evolution of the DM, baryonic, and cool gas masses. For instance, the DM mass enclosed in a given physical radius up to ∼75 kpc does not evolve between $z = 3.5$ and 2. The figure shows that DM growth is truncated from outside in consistent with a shut-down of accretion starting on large scales. Specifically, the DM mass with 50 kpc increases until at $z = 3.65$, while mass within 8 kpc grows until $z = 3.45$. The transition from a fast to a slow accretion phase also affects the evolution of the baryonic masses, the cool gas mass, and (omitted from Fig. 10 for reasons of presentational clarity) the stellar mass.

With the onset of this *'cosmological starvation'* phase the growth of the baryonic (and stellar mass) slows down substantially. It does not completely stop, however, because the accretion of cold gas

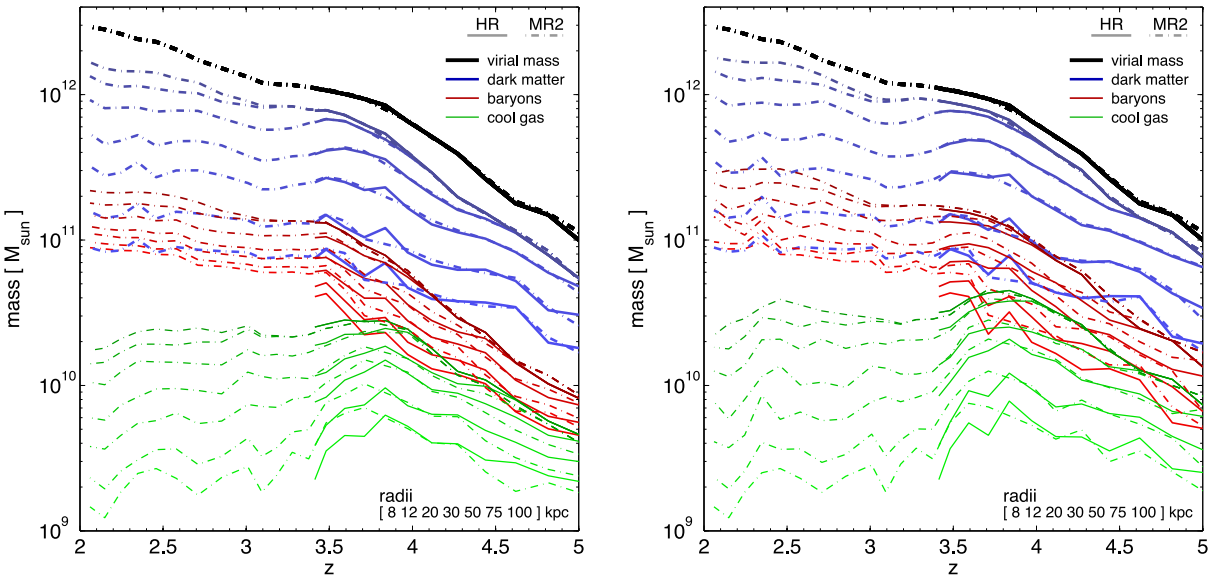

Figure 10. Evolution of DM and baryonic mass within fixed proper radii. Left-hand panel: only the contribution from the central galaxy and its halo are included. Satellite galaxies and subhaloes are removed. Right-hand panel: same as left-hand panel but including satellites and subhaloes. In each panel, red curves show the baryonic (stellar and gas) mass within fixed proper radii ranging from 8 kpc (lower line) to min(R_{vir}, 100 kpc) (upper line) as function of redshift. The blue lines show the evolution of the DM mass in the same fixed proper radii. The black lines indicate the virial mass of the halo. The virial mass increases continuously because of accretion at large radii ($\gtrsim$100 kpc) and the drop of the background density of the Universe with cosmic time. In contrast, DM and baryonic mass in *fixed proper radii* grow strongly only at $z > 3.5$. At $z = 2$–3.5, the growth of DM and baryons stalls.

from large radii supports a small, but non-negligible amount of star formation. In addition, mergers increase the stellar mass with *ex situ* formed stars. The mass of cool gas decreases slowly at $z < 3.7$, indicating that gas consumed in star formation is largely replenished by newly accreted gas. There is a marked difference in the cool gas-to-stellar mass ratio during the collapse phase ($\sim$1 at $z = 4$–5) and in the starvation phase ($<$0.1 at $z = 3.4$–2), see Table 2.

Cosmological starvation is not related to feedback from galaxies as we pointed out in Section 4. We prove this again in Fig. 11 which compares the evolution of the DM, stellar, and gas masses for the LR and LR-noFB runs. As expected, the no feedback run results in a somewhat larger stellar mass at any given time. However, aside from this change in the normalization the mass components evolve in the same way and show the same slow-down of their growth (or even decline) after $z \sim 3.5$.

We should point out that the virial mass continues to grow in the cosmological starvation phase. The increase in virial mass in the absence of physical accretion, sometimes called 'pseudo-evolution', is a consequence of tying the definition of virial mass to the evolving mean matter density (or the critical density) of the Universe (Diemand, Kuhlen & Madau 2007; Diemer, More & Kravtsov 2013). Hence, the virial mass does not robustly indicate whether a galaxy transits from the collapse phase to the starvation phase. Instead, the DM mass or the baryonic mass within fixed proper radii should be used.

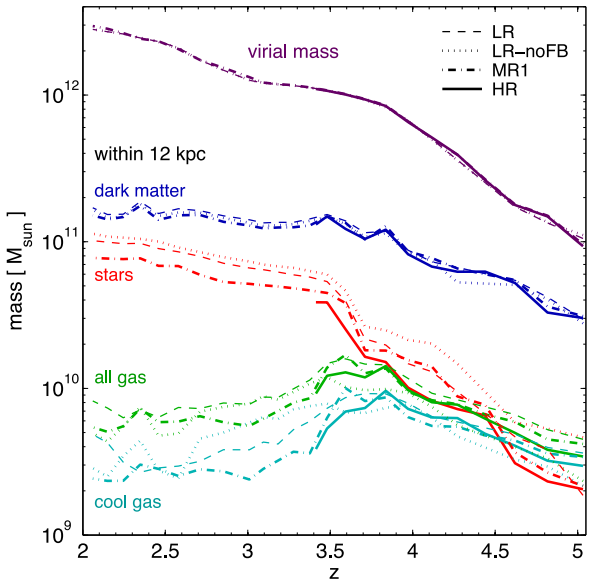

Figure 11. Evolution of DM and baryonic masses within a sphere of 12 proper kpc radius around the central galaxy and evolution of the virial mass. We show (from bottom to top) the mass of cool gas ($T < 2.5 \times 10^5$ K), the gas mass, the stellar mass, the DM mass (all within 12 kpc) as well as the virial mass of the halo (see legend). The dashed (dotted, dot–dashed, solid) lines show the mass components for the LR (LR-noFB, MR1, HR) run. The various mass components evolve in a similar way in the different runs suggesting that, on a qualitative level, our conclusions about cosmological starvation are neither affected by numerical resolution nor by the presence or absence of SN feedback.

6 CAVEATS

In this study, we analyse the evolution of a single massive galaxy at $z \sim 2$–5. The immediate question arises whether we can generalize our results based on a single object. In particular, it would be important to know whether our findings are representative of massive galaxies at high redshifts, or whether we accidentally picked a peculiar case. We selected the simulated galaxy based on its $z = 0$ halo mass and on the condition of residing in a 'typical' large-scale environment (based on the matter overdensity within 7 Mpc). Its halo accretion history differs from a purely exponential growth (Wechsler et al. 2002), as it grows quickly at early times and slows

down at later times. This is nothing special, though, as ∼40 per cent of haloes show such a behaviour (McBride, Fakhouri & Ma 2009). Furthermore, we show that the simulated galaxy is a reasonably good representation of quiescent massive galaxies observed at $z = 2$. Finally, we note that not all galaxies at high redshift are quiescent. In fact, roughly 20–50 per cent of massive galaxies are quiescent at $z = 2$ and the fraction decreases towards higher redshift (e.g. Muzzin et al. 2013). These results are consistent with a cosmological starvation origin of massive quiescent galaxies at high redshift, i.e. massive galaxies are quiescent at $z = 2$ if they reside in haloes with a particularly early halo formation time. Clearly, future simulations with larger galaxy samples will be helpful in challenging or confirming this picture.

By comparing the LR, MR and HR runs of the Argo simulation we can test whether the results are numerical converged. We find that global galaxy properties, such as stellar masses, gas masses, or gas inflow rates show only small trends with numerical resolution. The most noticeable differences are the sizes of galaxies at $z > 3$ and the entropy profile of hot gas around galaxies. The differences in sizes largely disappear by $z = 3$. Outside the very central region of the halo the entropy profiles of the HR and MR runs are in agreement. In contrast, the LR run produces hot circumgalactic gas with a significantly lower entropy, which reduces the gas cooling time and increases the star formation activity. Analogous trends have been noted before in cosmological simulations of disc galaxies (e.g. Mayer, Governato & Kaufmann 2008) and have been accredited to a reduced efficiency of SN feedback at low numerical resolution. Specifically, LR models struggle at capturing the stochastic, bursty, and inhomogeneous injection of thermal energy by SN explosions into the ISM thereby reducing the feedback efficiency. Overall, the agreement between the different runs is good (galaxy properties differ typically by less than 0.3 dex), especially at $z = 2$–3, indicating reasonable convergence.

The simulated galaxy leaves the star-forming sequence at $z \sim 3.5$ and reduces its SFR to a few $M_\odot\ yr^{-1}$ outside the central 300 pc (Fig. 7). In addition, the galaxy forms stars at a rate of $\sim$10 $M_\odot\ yr^{-1}$ in its very innermost (and numerical unresolved) region. Measuring the SFR of massive, quiescent (according to the *UVJ* classification) galaxies observed at high redshift is challenging (Hayward et al. 2014; Utomo et al. 2014). However, careful modelling of the spectral energy distributions of such galaxies shows that their SFRs are probably a few $M_\odot\ yr^{-1}$ or less (e.g. van de Sande et al. 2013, consistent with the non-central SFR of the simulated galaxy) although some 'quiescent galaxies' show SFRs up to $\sim$30 $M_\odot\ yr^{-1}$ (e.g. Straatman et al. 2014). Nonetheless, SFRs are probably less than a few $\sim M_\odot\ yr^{-1}$ in at least some observed high-redshift galaxies. It is unclear whether the physics implemented in our simulation is sufficient to reproduce massive galaxies with such low SFRs. We speculate that various AGN feedback channels (e.g. Sazonov et al. 2005; Croton et al. 2006) may play a role here acting *after* galaxies have already left the star-forming sequence. Possibly they are only effective after accretion of cool gas on to the galaxy has started to decline. Subsequently, they may provide a path to shut down star formation completely.

In this study, we took a minimalistic approach by including only physical processes that we understand reasonably well (hydrodynamics, gravity) or that we add to match observed properties of galaxies (SN feedback). However, we demonstrated that the choice of the feedback physics is not essential to explain why galaxies leave the star-forming sequence at high redshift. Specifically, we found no qualitative changes of our results when we turn off feedback processes altogether.

We should point out that the Argo simulation accounts for radiative cooling based on a primordial gas composition, but does not include cooling via metal lines. We justify this simplification of the physical realism of our simulations as follows. First, given the high temperatures (a few 10^6 K) and low metallicities ($Z \sim 0.1\ Z_\odot$) of the gas surrounding the simulated massive galaxy, we expect only moderate changes of the cooling rate (by about a factor of 2 or less, see below). Such a change will not affect our findings on a qualitative level in agreement with conclusions reached by previous work (e.g. Tescari et al. 2014). Secondly, uncertainties related to other aspects of our physical model (e.g. the modelling of feedback) largely outweigh the moderate quantitative changes that result from including or omitting metal line cooling. Thirdly, gas cooling is also affected by UV and soft X-ray emission from nearby massive stars (e.g. Cantalupo 2010; Kannan et al. 2014) and it is far from clear whether adding individual processes, e.g. metal line cooling, without their potential counterparts, e.g. radiative ionization by local sources, results in a better approximation of reality. Finally, we decided to mimic both the resolution and the physical model (which lacks high temperature metal line cooling) of the Eris simulation (Guedes et al. 2011) to extend its results to a higher mass regime.

In the left-hand panel of Fig. 12 we show the average gas temperature as function of distance from the galaxy at redshift 2, 3, and 4. The hot phase has a temperature of about 2–8 $\times\ 10^6$ K out to a large fraction of the virial radius. The metallicity of the circumgalactic medium is about 0.1 $Z_\odot$ at $z \geq 2$. Everything else being equal, metal line radiation increases the cooling rate by a factor of 2 at $T \sim 3.5 \times 10^6$ K compared with the cooling rate of gas with a primordial composition (Sutherland & Dopita 1993). However, adding metal line cooling in this straightforward way likely results in an overestimate of the cooling rate as we demonstrate now.

The right-hand panel of Fig. 12 shows the ionization parameter[4] as function of distance. The ionization parameter in the circumgalactic gas is ∼0.1–0.8 out to $R_{\rm vir}$ at $z = 2$–4. Photoionization by local sources (in addition to photoionization by a uniform UV background already included in our simulations) should both reduce the cooling efficiency of hot gas and heat the gas. An ionization parameter of ∼0.1 can suppress the net cooling by almost a factor of 2 at $T \sim 2 \times 10^6$ K (Cantalupo 2010) and by much larger factors at lower gas temperatures (e.g. Wiersma, Schaye & Smith 2009). In addition, ionization fractions may be out of equilibrium, which further reduces the cooling rate at $T < 10^6$ K (Gnat & Sternberg 2007; Oppenheimer & Schaye 2013). Finally, the radiation field of an AGN (not included in our simulation) could also heat the gas (via photoionization and Compton heating) and reduce the cooling (Ciotti & Ostriker 2007; Gnedin & Hollon 2012).

While our main findings should be robust against moderate changes in the cooling rates, we acknowledge the general need for a more sophisticated modelling of gas cooling in numerical simulations. In particular, local radiation fields (see Kannan et al. 2014 for a first study along these lines) and non-equilibrium effects (Oppenheimer & Schaye 2013) clearly matter, yet they are typically ignored in cosmological runs. A proper modelling of the cooling rates may also lower the amount of feedback required to offset excessive overcooling of gas.

[4] The ionization parameter is defined as the number density of ionizing photons divided by the hydrogen density. We compute it from equation 1 of Cantalupo (2010) using the SFR and the gas density profile of the simulated galaxy.

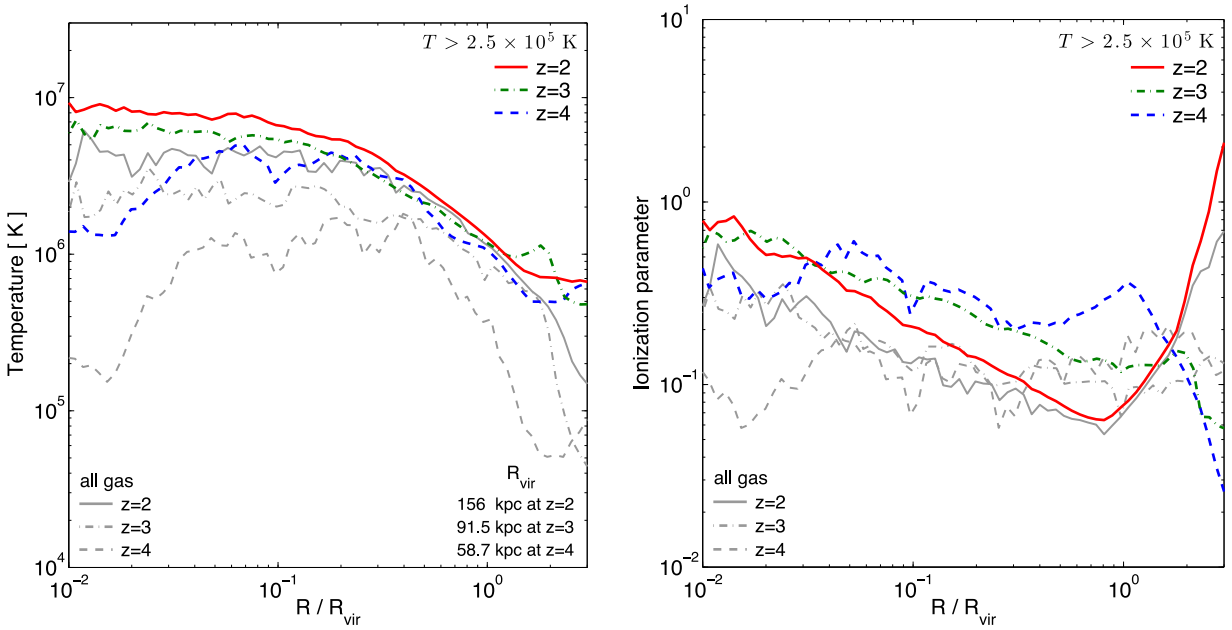

Figure 12. Left: average temperature versus distance from the centre of the galaxy. Right: ionization parameter versus distance from the centre of the galaxy. Plots are based on the MR2 run over the redshift range $z = 2$–4. The coloured lines show the temperature and the ionization parameter of the hot circumgalactic gas ($T > 2.5 \times 10^5$ K). Grey curves show the temperature and the ionization parameter if no temperature cut is imposed. The solid (dot–dashed, dashed) lines show the results for $z = 2$ ($z = 3$, 4). The hot circumgalactic gas has average temperatures of about ∼2–8 × 10^6 K out to a large fraction of the virial radius. The ionization parameter of the circumgalactic gas is 0.1–0.8.

7 DISCUSSION

In Section 5 we identified the levelling off and subsequent decrease of the gas accretion rate on to galaxies (cosmological starvation) as the trigger of star formation quenching in massive high-redshift galaxies. Based on our results we argued that major galaxy mergers and subsequent AGN-driven gas outflows, often suggested as the cause of star formation quenching (e.g. Springel et al. 2005b; Di Matteo et al. 2005), are likely of secondary importance. Such processes may however expedite the transition from the star forming to the quiescent population by reducing the gas reservoir of galaxies via central starbursts and strong AGN-driven outflows. However, as Fig. 7 shows, the suppression of the SFR can be quite fast ($\tau \sim 100$ Myr) even in the case without AGN feedback or the occurrence of a major merger.

The process that quenched star formation and let to quiescent, early-type galaxies in the local Universe probably lasted no longer than a few hundred Myr (e.g. Schawinski et al. 2014). In contrast, the typical gas depletion time of a 'normal' star forming, nearby spiral is a few Gyr (e.g. Bigiel et al. 2008). The difference in time-scales has been used to argue that AGN feedback is thus required to suppress star formation on sufficiently fast time-scales (Kaviraj et al. 2011). Here we demonstrate, however, that such a rapid decline in the SFR is also possible in compact, massive galaxies (which have short dynamical times) with the help of SN feedback.

A different time-scale constraint can be derived from the abundance ratio of α elements and iron in stellar atmospheres. It is typically understood that this ratio measures the time-scale of star formation. Unfortunately, the observed ratio is also sensitive to the IMF, to the time delay function of Type Ia SNe, to the adopted stellar yields, and to metal losses via stellar winds, which complicates its interpretation. Star formation time-scales derived this way typically range from ∼0.2–0.8 Gyr (see Conroy, Graves & van Dokkum 2014, and references therein), up to ∼1 Gyr (Thomas, Greggio & Bender 1999), but are broadly consistent with our findings (few hundred Myr).

We showed in Fig. 8 that the gas accretion rate shuts down before the SFR. Hence, even after cosmological starvation sets in, galaxies may continue their star formation and AGN activity for up to a few ∼100 Myr. If feedback from the AGN does not affect the dynamics of the ISM, galaxies likely shut down their activity from outside in, i.e. a declining SFR should precede a declining AGN activity. In this case, galaxies on the star-forming sequence as well as recently quenched galaxies may show AGN activity.

An idea often associated with the suppression of star formation in massive haloes is the concept of a threshold halo mass (Birnboim & Dekel 2003; Cattaneo et al. 2006; Dekel & Birnboim 2006; Bouché et al. 2010). The threshold mass arises from the competition of gas cooling and gas compression in the shock-heated circumgalactic medium. The cooling time in massive haloes is sufficiently long to support a stable virial shock that heats most of the accreted gas to the virial temperature of the halo. Additional heating sources such as radio mode AGN feedback or gravitational heating may further reduce the already low cooling rates of this shock-heated circumgalactic material.

However, numerical simulations show that at $z \geq 2$ dense filaments of cool gas can penetrate the virial radius of even massive haloes unshocked (Kereš et al. 2005, 2009; Ocvirk, Pichon & Teyssier 2008; Nelson et al. 2013). In fact, much of the star formation activity at high redshift is likely supported by inflowing streams of cool gas (Dekel et al. 2009). Hence, the threshold mass picture may explain the origin of massive, quiescent galaxies at $z < 2$ (Cattaneo et al. 2006), but it does not account for the large number of quiescent galaxies observed at $z \geq 2$. Cosmological starvation is a complementary process that solves this problem by reducing the large-scale gas accretion rate on to selected haloes at high redshift. It is primarily a gravitational, not a hydrodynamical,

process and affects all massive, high-redshift galaxies that reside in haloes undergoing a period of reduced growth.

Cosmological starvation also differs from the two-phase model of galaxy formation developed to explain the size growth of massive galaxies (Naab et al. 2007; Naab, Johansson & Ostriker 2009; Feldmann et al. 2010; Oser et al. 2010, 2012). The two phase model consists of an early dissipative phase in which much of the stellar mass of the central galaxy is formed resulting in a compact massive galaxy at high redshift. During the second phase of the two-phase model the (by then) quiescent galaxy grows in mass and size by mergers, i.e. via the accretion of stellar material. In contrast, the cosmological starvation phase is a period marked by little accretion of both gaseous and stellar matter. It is responsible for the abundance of massive, compact, quiescent galaxies at high redshift.

The picture we propose in this work is consistent with several previous findings. *N*-body simulations of DM often show a transition from a period of genuine growth to a period of starvation (e.g. Diemand et al. 2007). During starvation the mass within fixed proper radii remains constant, while the virial mass still grows ('pseudo-evolution'; Diemer et al. 2013). However, the importance of the starvation phase for the evolution of galaxies has not yet been fully appreciated, despite the close correlation between DM accretion and gas accretion on to haloes (Faucher-Giguère, Kereš & Ma 2011; van de Voort et al. 2011).

Empirical models that correlate colours and star formation activity of galaxies at fixed stellar mass with the *halo formation time* are successful in predicting the clustering of galaxies as a function of their colour or star formation activity, at least for galaxies in the local Universe (Hearin & Watson 2013; Hearin et al. 2014; Watson et al. 2014). Cosmological starvation (potentially combined with the formation of stable virial shocks) is a promising candidate that may provide the physical basis for the main ansatz of these models.

Models that attribute quenching to feedback, from stellar sources or via an AGN, face the challenge to explain why the quiescent and the star-forming galaxy populations substantially overlap in stellar mass (at 10^{11} $M_\odot$) at high redshift (e.g. Magnelli et al. 2014). If star formation quenching is related to feedback from the galaxy, why do some galaxies quench and other, similarly massive galaxies, do not? This observation is also a challenge for models based solely on a halo mass threshold. In such models one expects quenching to occur as soon as galaxies exceed a critical stellar mass. In contrast, the cosmological starvation picture naturally accounts for the co-existence of star-forming and quiescent galaxies of the same stellar (and halo) mass. Here, the star formation activity is based on the recent rate of gas accretion rate on to the halo and not just on the halo mass.

We can only speculate why the importance of cosmological starvation has not been recognized before. First, our work is probably the first cosmological simulation that reproduces the observed global properties of a massive, high-redshift galaxy. Previous simulations with insufficient resolution and/or inadequate physics might not have recognized the starvation phase because of overcooling. Also, excessively strong feedback could have masked the modulation of gas accretion. Secondly, previous work often focused on *average* accretion rates and SFRs for haloes of a given mass (e.g. Faucher-Giguère et al. 2011; van de Voort et al. 2011). However, the point of the starvation model is that it distinguishes between haloes of the same mass but with a different accretion history. Also, the fraction of massive, quiescent galaxies is only ∼20 per cent at $z \gtrsim 3$ (Muzzin et al. 2013). Hence, the impact of this significant galaxy population can be missed if only average or median properties are studied. Furthermore, the pseudo-evolution of the halo mass (Diemer et al. 2013) may obfuscate the connection between star formation and DM accretion. Pseudo-evolution has so far been ignored in galaxy evolution studies based on DM merger trees. Finally, as many simulations still struggle in producing galaxies with realistic stellar masses and SFRs, less attention has been given to study exactly when and how galaxies leave the star-forming sequence.

8 SUMMARY AND CONCLUSIONS

One of the major unsolved problems in galaxy evolution is to identify the physical process, or the processes, regulating the transition of galaxies from the star forming to the quiescent population. The old ages of massive, quiescent galaxies in the local Universe place this quenching of star formation at an early time, likely before $z \sim 2$ (Thomas et al. 2010). Similarly, the ages of massive, quiescent galaxies at $z = 2$–4 (e.g. Whitaker et al. 2013; Straatman et al. 2014) indicate that some galaxies reduce their SFRs to low levels at even higher redshift. In this paper, we analyse the origin of star formation quenching in high-redshift galaxies with the help of a state-of-the-art cosmological simulation of a massive galaxy.

Our main findings are as follows.

(i) The global properties of the simulated galaxy are in good agreement with those of similarly massive galaxies observed at high redshift. Reproduced properties include the stellar-to-virial mass ratio, the size of the stellar component, and the sSFR while on the star-forming sequence.

(ii) The simulation includes thermal SN feedback, but does not model feedback from AGN. This suggests that AGN feedback is not an essential ingredient to reproduce properties of massive, quiescent galaxies at high redshift.

(iii) At $z \sim 3.5$ the simulated galaxy leaves the star-forming sequence and the sSFRs decrease by almost an order of magnitude within a few 100 Myr. The SFR declines approximately exponentially with an e-fold time of ∼ 100 Myr. By $z \sim 2$ the colours and the stellar mass of the simulated galaxy agree with those of massive, quiescent galaxies present at those redshifts.

(iv) The drop of the sSFR is not caused by feedback processes, but rather a consequence of a levelling off and subsequent decline in the cool gas accretion rate on to the halo of the galaxy.

(v) The decrease of the sSFR is somewhat faster and more pronounced in our higher resolution runs compared with the LR run. We attribute this difference to the diminished efficacy of the implemented stellar feedback scheme at low numerical resolution. Hence, feedback likely plays a crucial role in suppressing star formation to the very low levels (less than a few $M_\odot$ yr^{-1}) observed in a large fraction of massive, quiescent, high-redshift galaxies.

Based on our findings we propose a novel picture for the suppression of star formation that differs from previous suggestions based on a halo mass threshold or on star formation quenching via feedback-driven outflows during major mergers. After a period of fast gas accretion and exponential growing SFRs, some massive galaxies at high redshift enter a period of cosmological starvation in which the gas and DM accretion rates on to their haloes first stall and subsequently decrease. Affected galaxies leave the main sequence as the stalled gas accretion rates no longer support exponentially increasing SFRs.

The cosmological starvation picture physically connects the sSFR of high-redshift galaxies to the gas accretion rate on to their haloes. It makes a number of testable predictions.

Generally, we expect that central galaxies residing in unrelaxed, still collapsing large-scale structures have larger sSFRs than central

galaxies embedded in a relaxed, virialized environment. Specifically, we predict that at $z \geq 2$ central galaxies of a given stellar mass are more likely to be still star forming if they reside in a higher density environment, e.g. are surrounded by a larger number of satellite galaxies. In fact, this reversal of the local star forming–density relation, namely larger sSFRs in denser regions, has been observed at $z \sim 1$ (Elbaz et al. 2007).

Also, we expect the region of shock-heated gas to expand outward during a period of reduced gas accretion rate and, thus, lower density of the pre-shocked infalling gas (see equation 29 in Birnboim & Dekel 2003). Hence, cosmological starvation might also be tested by inferring and comparing the size of the virial shock around massive, quiescent galaxies and around star-forming galaxies of the same mass. At $z < 1$ environmental processes related to the presence of a hot, dilute atmosphere of shock-heated gas affect satellite galaxies and help in suppressing their SFR (e.g. Kovač et al. 2014). Satellite galaxies around massive, high-redshift galaxies may thus potentially be used to probe the extent of the virial shock.

Finally, during the starvation phase the accretion of both gas and stellar material is reduced. Hence, we predict that recently quenched, high-redshift galaxies have (on average) a different distribution of satellite galaxies (e.g. fewer satellites at large distances, fewer massive satellites) than star-forming galaxies of the same stellar mass.

Cosmological starvation highlights the importance of the recent halo accretion history for the evolution of galaxies. Abundance matching techniques show that the stellar mass of galaxies is primarily controlled by their halo mass. Here we demonstrate that the accretion rate on to haloes acts as an additional lever that controls the SFR of central galaxies at high redshift. Accretion rates and masses are tightly coupled for haloes with an average accretion history (Neistein & Dekel 2008), but differ for the large fraction of haloes that deviate from pure exponential growth (McBride et al. 2009). The exploration of the large variety of halo accretion histories in future work may lead to a deeper understanding of the observed diversity of high-redshift galaxies.

ACKNOWLEDGEMENTS

We thank A. Babul, G. Barro, M. Kriek, P. Madau, D. Martizzi, D. Marchesini, J. X. Prochaska, E. Quataert, E. Scannapieco, and F. van de Voort for inspiring discussions and valuable input. We are also grateful to the Aspen Center for Physics for organizing the winter 2014 conference 'Unveiling the Formation of Massive Galaxies – Theoretical and Observational Challenges', which gave us the opportunity to present the results of this work to, and to receive stimulating suggestions from, the galaxy evolution community. We thank the Fermi National Accelerator Laboratory for supercomputing time. We also thank the Swiss Supercomputing Center for granting us early user access to the new Cray XC30 'Piz Daint' on which the HR run is being carried out. RF acknowledges support for this work by NASA through Hubble Fellowship grant HF-51304.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. This work made extensive use of the NASA Astrophysics Data System and arXiv.org preprint server.

REFERENCES

Abadi M. G., Moore B., Bower R. G., 1999, MNRAS, 308, 947
Alexander D. M., Hickox R. C., 2012, New Astron. Rev., 56, 93
Arnouts S. et al., 2007, A&A, 476, 137
Aumer M., White S. D. M., Naab T., Scannapieco C., 2013, MNRAS, 434, 3142
Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017
Balogh M. L., Navarro J. F., Morris S. L., 2000, ApJ, 540, 113
Balogh M. L., Baldry I. K., Nichol R., Miller C., Bower R., Glazebrook K., 2004, ApJ, 615, L101
Barro G. et al., 2013, ApJ, 765, 104
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Bell E. F. et al., 2004, ApJ, 608, 752
Bertschinger E., 1995, preprint (astro-ph/9506070)
Bertschinger E., 2001, ApJS, 137, 1
Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25
Bezanson R., van Dokkum P., van de Sande J., Franx M., Kriek M., 2013, ApJ, 764, L8
Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2846
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Blanton M. R. et al., 2003, ApJ, 594, 186
Borgani S., Kravtsov A., 2011, Adv. Sci. Lett., 4, 204
Bouché N. et al., 2010, ApJ, 718, 1001
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Boyle B. J., Terlevich R. J., 1998, MNRAS, 293, L49
Brammer G. B. et al., 2009, ApJ, 706, L173
Brammer G. B. et al., 2011, ApJ, 739, 24
Brook C. B., Stinson G., Gibson B. K., Wadsley J., Quinn T., 2012, MNRAS, 424, 1275
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cantalupo S., 2010, MNRAS, 403, L16
Carollo C. M. et al., 2013, ApJ, 773, 112
Cassata P. et al., 2011, ApJ, 743, 96
Cattaneo A., Dekel A., Devriendt J., Guiderdoni B., Blaizot J., 2006, MNRAS, 370, 1651
Cen R., 2014, ApJ, 781, 38
Chabrier G., 2003, PASP, 115, 763
Cimatti A. et al., 2004, Nature, 430, 184
Cimatti A. et al., 2008, A&A, 482, 21
Ciotti L., Ostriker J. P., 2007, ApJ, 665, 1038
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Conroy C., Graves G. J., van Dokkum P. G., 2014, ApJ, 780, 33
Croton D. J. et al., 2006, MNRAS, 365, 11
Daddi E. et al., 2005, ApJ, 626, 680
Daddi E. et al., 2007, ApJ, 670, 156
Daddi E., Dannerbauer H., Elbaz D., Dickinson M., Morrison G., Stern D., Ravindranath S., 2008, ApJ, 673, L21
Daddi E. et al., 2010, ApJ, 713, 686
Damjanov I. et al., 2011, ApJ, 739, L44
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
Debuhr J., Quataert E., Ma C.-P., Hopkins P., 2010, MNRAS, 406, L55
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A., Birnboim Y., 2008, MNRAS, 383, 119
Dekel A., Silk J., 1986, ApJ, 303, 39
Dekel A. et al., 2009, Nature, 457, 451
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Diemand J., Kuhlen M., Madau P., 2007, ApJ, 667, 859
Diemer B., More S., Kravtsov A. V., 2013, ApJ, 766, 25
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2010, MNRAS, 409, 985
Dubois Y., Devriendt J., Teyssier R., Slyz A., 2011, MNRAS, 417, 1853
Dubois Y., Gavazzi R., Peirani S., Silk J., 2013, MNRAS, 433, 3297
Elbaz D. et al., 2007, A&A, 468, 33
Elbaz D. et al., 2011, A&A, 533, A119

Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, ApJ, 644, 813
Farouki R., Shapiro S. L., 1981, ApJ, 243, 32
Faucher-Giguère C.-A., Quataert E., 2012, MNRAS, 425, 605
Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982
Feldmann R., 2013, MNRAS, 433, 1910
Feldmann R., Carollo C. M., Mayer L., Renzini A., Lake G., Quinn T., Stinson G. S., Yepes G., 2010, ApJ, 709, 218
Feldmann R., Carollo C. M., Mayer L., 2011, ApJ, 736, 88
Finlator K., Oppenheimer B. D., Davé R., 2011, MNRAS, 410, 1703
Fontana A. et al., 2009, A&A, 501, 15
Förster Schreiber N. M. et al., 2004, ApJ, 616, 40
Förster Schreiber N. M. et al., 2006, ApJ, 645, 1062
Franx M. et al., 2003, ApJ, 587, L79
Franx M., van Dokkum P. G., Schreiber N. M. F., Wuyts S., Labbé I., Toft S., 2008, ApJ, 688, 770
Fumagalli M. et al., 2014, ApJ, 796, 35
Gabor J. M., Bournaud F., 2014, MNRAS, 441, 1615
Gill S. P. D., Knebe A., Gibson B. K., 2004, MNRAS, 351, 399
Gnat O., Sternberg A., 2007, ApJS, 168, 213
Gnedin N. Y., Hollon N., 2012, ApJS, 202, 13
Goulding A. D. et al., 2014, ApJ, 783, 40
Governato F. et al., 2010, Nature, 463, 203
Guedes J., Callegari S., Madau P., Mayer L., 2011, ApJ, 742, 76
Gunn J. E., Gott J. R., III, 1972, ApJ, 176, 1
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Haardt F., Madau P., 1996, ApJ, 461, 20
Haas M. R., Schaye J., Booth C. M., Dalla Vecchia C., Springel V., Theuns T., Wiersma R. P. C., 2013, MNRAS, 435, 2931
Hahn O., Porciani C., Carollo C. M., Dekel A., 2007, MNRAS, 375, 489
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hayward C. C. et al., 2014, MNRAS, 445, 1598
Hearin A. P., Watson D. F., 2013, MNRAS, 435, 1313
Hearin A. P., Watson D. F., Becker M. R., Reyes R., Berlind A. A., Zentner A. R., 2014, MNRAS, 444, 729
Hewett P. C., Warren S. J., Leggett S. K., Hodgkin S. T., 2006, MNRAS, 367, 454
Hogg D. W. et al., 2004, ApJ, 601, L29
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1
Hopkins P. F., Keres D., Onorbe J., Faucher-Giguere C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Humphrey P. J., Buote D. A., Canizares C. R., Fabian A. C., Miller J. M., 2011, ApJ, 729, 53
Ilbert O. et al., 2010, ApJ, 709, 644
Johansson P. H., Naab T., Ostriker J. P., 2009, ApJ, 697, L38
Johansson P. H., Naab T., Ostriker J. P., 2012, ApJ, 754, 115
Kannan R. et al., 2014, MNRAS, 437, 2882
Kauffmann G. et al., 2003a, MNRAS, 341, 33
Kauffmann G. et al., 2003b, MNRAS, 346, 1055
Kaviraj S., Schawinski K., Silk J., Shabala S. S., 2011, MNRAS, 415, 3798
Kawata D., Mulchaey J. S., 2008, ApJ, 672, L103
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Khalatyan A., Cattaneo A., Schramm M., Gottlöber S., Steinmetz M., Wisotzki L., 2008, MNRAS, 387, 13
Khochfar S., Ostriker J. P., 2008, ApJ, 680, 54
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kocevski D. D. et al., 2012, ApJ, 744, 148
Kovač K. et al., 2014, MNRAS, 438, 717
Kriek M., van der Wel A., van Dokkum P. G., Franx M., Illingworth G. D., 2008, ApJ, 682, 896
Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545
Labbé I. et al., 2005, ApJ, 624, L81
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692
Le Brun A. M. C., McCarthy I. G., Schaye J., Ponman T. J., 2014, MNRAS, 441, 1270
Lee K.-S. et al., 2011, ApJ, 733, 99
Lundgren B. F. et al., 2014, ApJ, 780, 34
McBride J., Fakhouri O., Ma C.-P., 2009, MNRAS, 398, 1858
McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593
McCarthy I. G. et al., 2010, MNRAS, 406, 822
McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, MNRAS, 412, 1965
McCracken H. J. et al., 2010, ApJ, 708, 202
McKee C. F., Ostriker J. P., 1977, ApJ, 218, 148
Magdis G. E. et al., 2012, ApJ, 760, 6
Magnelli B. et al., 2014, A&A, 561, A86
Magorrian J. et al., 1998, AJ, 115, 2285
Maiolino R. et al., 2008, A&A, 488, 463
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115
Marchesini D. et al., 2014, ApJ, 794, 65
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Martizzi D., Teyssier R., Moore B., 2012, MNRAS, 420, 2859
Mayer L., Governato F., Kaufmann T., 2008, Adv. Sci. Lett., 1, 7
Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613
More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, MNRAS, 392, 801
Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Munshi F. et al., 2013, ApJ, 766, 56
Murray N., Quataert E., Thompson T. A., 2005, ApJ, 618, 569
Muzzin A. et al., 2013, ApJ, 777, 18
Naab T., Johansson P. H., Ostriker J. P., Efstathiou G., 2007, ApJ, 658, 710
Naab T., Johansson P. H., Ostriker J. P., 2009, ApJ, 699, L178
Neistein E., Dekel A., 2008, MNRAS, 383, 615
Nelson D., Vogelsberger M., Genel S., Sijacki D., Kereš D., Springel V., Hernquist L., 2013, MNRAS, 429, 3353
Newman A. B., Ellis R. S., Treu T., Bundy K., 2010, ApJ, 717, L103
Newton R. D. A., Kay S. T., 2013, MNRAS, 434, 3606
Noeske K. G. et al., 2007, ApJ, 660, L43
Novak G. S., Ostriker J. P., Ciotti L., 2011, ApJ, 737, 26
Ocvirk P., Pichon C., Teyssier R., 2008, MNRAS, 390, 1326
Onodera M. et al., 2012, ApJ, 755, 26
Oppenheimer B. D., Schaye J., 2013, MNRAS, 434, 1043
Oser L., Ostriker J. P., Naab T., Johansson P. H., Burkert A., 2010, ApJ, 725, 2312
Oser L., Naab T., Ostriker J. P., Johansson P. H., 2012, ApJ, 744, 63
Ostriker J. P., Choi E., Ciotti L., Novak G. S., Proga D., 2010, ApJ, 722, 642
Patel S. G. et al., 2013, ApJ, 766, 15
Peng Y.-j. et al., 2010, ApJ, 721, 193
Peng Y.-j., Lilly S. J., Renzini A., Carollo M., 2012, ApJ, 757, 4
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Power C., Nayakshin S., King A., 2011, MNRAS, 412, 269
Puchwein E., Sijacki D., Springel V., 2008, ApJ, 687, L53
Ragone-Figueroa C., Granato G. L., Murante G., Borgani S., Cui W., 2013, MNRAS, 436, 1750
Raiteri C. M., Villata M., Navarro J. F., 1996, A&A, 315, 105
Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30
Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74
Sargent M. T. et al., 2014, ApJ, 793, 19
Sazonov S. Y., Ostriker J. P., Ciotti L., Sunyaev R. A., 2005, MNRAS, 358, 168
Scannapieco E., Oh S. P., 2004, ApJ, 608, 62
Scannapieco C. et al., 2012, MNRAS, 423, 1726
Schawinski K., Treister E., Urry C. M., Cardamone C. N., Simmons B., Yi S. K., 2011, ApJ, 727, L31
Schawinski K. et al., 2014, MNRAS, 440, 889

Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978
Shen S., Madau P., Guedes J., Mayer L., Prochaska J. X., Wadsley J., 2013, ApJ, 765, 89
Shen S., Madau P., Conroy C., Governato F., Mayer L., 2014, ApJ, 792, 99
Sijacki D., Springel V., 2006, MNRAS, 366, 397
Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877
Spergel D. N. et al., 2007, ApJS, 170, 377
Springel V., Di Matteo T., Hernquist L., 2005a, MNRAS, 361, 776
Springel V., Di Matteo T., Hernquist L., 2005b, ApJ, 620, L79
Stadel J. G., 2001, PhD thesis, Univ. Washington
Stinson G., Seth A., Katz N., Wadsley J., Governato F., Quinn T., 2006, MNRAS, 373, 1074
Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, MNRAS, 428, 129
Straatman C. M. S. et al., 2014, ApJ, 783, L14
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Szomoru D., Franx M., van Dokkum P. G., 2012, ApJ, 749, 121
Tacconi L. J. et al., 2010, Nature, 463, 781
Tacconi L. J. et al., 2013, ApJ, 768, 74
Tal T. et al., 2014, ApJ, 789, 164
Taylor E. N. et al., 2009, ApJ, 694, 1171
Tescari E., Katsianis A., Wyithe J. S. B., Dolag K., Tornatore L., Barai P., Viel M., Borgani S., 2014, MNRAS, 438, 3490
Teyssier R., Moore B., Martizzi D., Dubois Y., Mayer L., 2011, MNRAS, 414, 195
Thomas D., Greggio L., Bender R., 1999, MNRAS, 302, 537
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Toft S. et al., 2007, ApJ, 671, 285
Toft S., Gallazzi A., Zirm A., Wold M., Zibetti S., Grillo C., Man A., 2012, ApJ, 754, 3
Treister E., Schawinski K., Urry C. M., Simmons B. D., 2012, ApJ, 758, L39
Trujillo I. et al., 2006, ApJ, 650, 18
Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16
Utomo D., Kriek M., Labbé I., Conroy C., Fumagalli M., 2014, ApJ, 783, L30
Vale A., Ostriker J. P., 2004, MNRAS, 353, 189
van de Sande J. et al., 2011, ApJ, 736, L9
van de Sande J. et al., 2013, ApJ, 771, 85
van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, MNRAS, 414, 2458
van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79
van der Wel A., Holden B. P., Zirm A. W., Franx M., Rettura A., Illingworth G. D., Ford H. C., 2008, ApJ, 688, 48
van Dokkum P. G. et al., 2006, ApJ, 638, L59
van Dokkum P. G. et al., 2008, ApJ, 677, L5
van Dokkum P. G., Kriek M., Franx M., 2009, Nature, 460, 717
van Dokkum P. G. et al., 2010, ApJ, 709, 1018
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Voit G. M., Bryan G. L., 2001, Nature, 414, 425
Wadsley J. W., Stadel J., Quinn T., 2004, New Astron., 9, 137
Watson D. F., Hearin A. P., Berlind A. A., Becker M. R., Behroozi P. S., Skibba R. A., Reyes R., Zentner A. R., 2014, preprint (arXiv:1403.1578)
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, MNRAS, 432, 336
Whitaker K. E. et al., 2011, ApJ, 735, 86
Whitaker K. E. et al., 2013, ApJ, 770, L39
Wiersma R. P. C., Schaye J., Smith B. D., 2009, MNRAS, 393, 99
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Toft S., Kriek M., Labbé I., 2010, ApJ, 713, 738
Wurster J., Thacker R. J., 2013, MNRAS, 431, 2513
Wuyts S. et al., 2007, ApJ, 655, 51
Wuyts S. et al., 2011, ApJ, 742, 96

APPENDIX A: FUNCTIONAL FORM OF THE STAR FORMATION HISTORY

We can motivate exponentially growing and declining star formation histories (equation 1) based on the conservation of gas mass and the ansatz of a linear relation between SFR and gas mass. The latter ansatz is not strictly accurate for the Argo simulation, as it uses a non-linear relation between SFR density and gas density, see Section 2. In fact, this non-linear relation implies that the SFR depends not only on gas mass, but also on the density distribution of the gas. The following 'derivation' of equation (1) ignores these complications.

The mass of gas, $M_{\rm g}$, available for star formation in a galaxy changes because of gas accretion ($\dot{M}_{\rm g,acc}$), star formation (SFR $= M_{\rm g}/\tau_{\rm dep}$), galactic outflows (mass loading factor $\epsilon_{\rm out}$), and stellar mass-loss (return fraction R), i.e.

$$\dot{M}_{\rm gas} = \dot{M}_{\rm gas,acc} - \frac{1-R+\epsilon_{\rm out}}{\tau_{\rm dep}} M_{\rm gas}. \tag{A1}$$

We can easily solve this differential equation if we assume that $\hat{\tau}_{\rm dep} = \frac{\tau_{\rm dep}}{1-R+\epsilon_{\rm out}}$ is a constant, and that $\dot{M}_{\rm gas,acc}$ depends on time, but not on $M_{\rm gas}$ itself. Basic calculus shows that the differential equation $\dot{x} + \beta x = \gamma(t)$ has the general solution $x(t) = (\Gamma(t) + C)e^{-\beta t}$ with $\Gamma(t) = \int^t \gamma(t')e^{\beta t'}{\rm d}t'$ and constant C.

We now assume that there are two phases in the life of a galaxy. A collapse phase in which gas accretion increases approximately exponentially with time t since the big bang, i.e. $\dot{M}_{\rm g,acc}(t) = \gamma_{\rm G}\, e^{t/\tau_{\rm acc}}$ and a cosmological starvation phase in which the gas accretion levels off ($\dot{M}_{\rm g,acc} = $ const.) or stops altogether ($\dot{M}_{\rm g,acc} \approx 0$). We further assume that the gas mass is zero at time $t = 0$.

During the collapse phase the gas mass at time t is

$$\begin{aligned} M_{\rm gas}(t) &= \frac{\gamma_{\rm G}}{\tau_{\rm acc}^{-1} + \hat{\tau}_{\rm dep}^{-1}} \left(e^{t/\tau_{\rm acc}} - e^{-t/\hat{\tau}_{\rm dep}}\right) \\ &\approx \frac{\gamma_{\rm G}}{\tau_{\rm acc}^{-1} + \hat{\tau}_{\rm dep}^{-1}} e^{t/\tau_{\rm acc}} \text{ if } t \gtrsim \tau_{\rm acc}. \end{aligned}$$

Hence,

$$\text{SFR} \approx \frac{\gamma_{\rm G}}{\tau_{\rm dep}/\tau_{\rm acc} + 1 - R + \epsilon_{\rm out}} e^{t/\tau_{\rm acc}} = \dot{M}_{\rm g,acc}\, \mathcal{O}(1),$$

i.e. the SFR mirrors the gas accretion rate. The SFR is a pure exponential with zero offset during this phase.

For the starvation phase and the initial conditions $\text{SFR}(t^*) = \text{SFR}^*$, we obtain

$$\text{SFR} = O + (\text{SFR}^* - O)e^{-(t-t^*)/\hat{\tau}_{\rm dep}}, \text{ with } O = \frac{\dot{M}_{\rm g,acc}}{1-R+\epsilon_{\rm out}}.$$

Here t^* is the time when the galaxy changes from the collapse phase into the starvation phase, i.e. the time when it leaves the star-forming sequence. Physically, the SFR declines in the starvation phase on a gas depletion time-scale because the gas reservoir is used up by star formation. The offset from a purely exponential decline is a consequence of continued gas accretion.

This paper has been typeset from a TEX/LATEX file prepared by the author.